

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 23, 2003, 19:50:48 ; Search time 159 Seconds  
(without alignments)  
2058.090 Million cell updates/sec

Title: us-09-817-199b-2

Perfect score: 1150

Sequence: 1 MWTPCAVATRGEAPERSP.....FQIRDVFSQKKRSCCSFPM 223

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-framer\_p2n.model -DRV-xlh  
-Q-/cgn2\_1/USPTO.spool/US09817199/runat\_18062003\_145103\_857/app\_query.fasta\_1.391  
-DB-PublishedApplications\_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62  
-TRANS-human40.cd1 -LISP=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09817199 -SCGN\_1\_57\_@runat\_18062003\_145103\_857  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                         |
|------------|-------|-------------|--------|----|-------------------------------------|
| 1          | 1150  | 100.0       | 1106   | 9  | US-10-037-270-959 Sequence 959, App |
| 2          | 1150  | 100.0       | 2674   | 10 | US-09-817-199A-1 Sequence 1, Appl   |
| 3          | 1145  | 99.6        | 1116   | 10 | US-09-794-257-13 Sequence 13, Appl  |
| 4          | 1144  | 99.5        | 2623   | 9  | US-09-764-868-71 Sequence 71, Appl  |

|    |       |      |      |    |                     |                   |
|----|-------|------|------|----|---------------------|-------------------|
| 5  | 1140  | 99.1 | 875  | 12 | US-10-051-986-10    | Sequence 10, Appl |
| 6  | 1133  | 98.5 | 1316 | 9  | US-09-764-868-493   | Sequence 493, App |
| 7  | 977   | 85.0 | 576  | 10 | US-09-794-257-15    | Sequence 15, Appl |
| 8  | 720   | 62.6 | 447  | 10 | US-09-867-550-1811  | Sequence 1811, Ap |
| 9  | 540   | 47.0 | 434  | 9  | US-09-918-995-648   | Sequence 648, App |
| 10 | 535   | 46.5 | 964  | 9  | US-09-764-868-75    | Sequence 75, Appl |
| 11 | 511.5 | 44.5 | 2497 | 10 | US-09-834-975-879   | Sequence 879, App |
| 12 | 511.5 | 44.5 | 2497 | 10 | US-09-834-975-885   | Sequence 885, App |
| 13 | 511.5 | 44.5 | 2497 | 10 | US-09-834-975-894   | Sequence 894, App |
| 14 | 511.5 | 44.5 | 2497 | 10 | US-09-834-975-896   | Sequence 896, App |
| 15 | 507   | 44.1 | 881  | 10 | US-09-770-445-529   | Sequence 529, App |
| 16 | 506   | 44.0 | 1537 | 10 | US-09-925-300-631   | Sequence 631, App |
| 17 | 501   | 43.6 | 624  | 10 | US-09-794-257-9     | Sequence 9, Appl  |
| 18 | 501   | 43.6 | 1161 | 10 | US-09-794-257-7     | Sequence 7, Appl  |
| 19 | 500   | 43.5 | 651  | 9  | US-09-938-842A-836  | Sequence 836, App |
| 20 | 496.5 | 43.2 | 609  | 9  | US-09-938-842A-832  | Sequence 832, App |
| 21 | 483   | 42.0 | 1274 | 10 | US-09-925-302-91    | Sequence 91, Appl |
| 22 | 470   | 40.9 | 925  | 10 | US-09-967-736-4     | Sequence 4, Appl  |
| 23 | 469.5 | 40.8 | 939  | 9  | US-10-102-806-48    | Sequence 48, Appl |
| 24 | 448.5 | 39.0 | 639  | 10 | US-09-350-874-66    | Sequence 66, Appl |
| 25 | 446   | 38.8 | 833  | 9  | US-09-764-868-88    | Sequence 88, Appl |
| 26 | 435   | 37.8 | 2021 | 9  | US-09-817-198A-1    | Sequence 1, Appl  |
| 27 | 435   | 37.6 | 3257 | 10 | US-09-764-868-488   | Sequence 488, App |
| 28 | 432   | 37.6 | 1000 | 9  | US-09-764-868-507   | Sequence 507, App |
| 29 | 429.5 | 37.3 | 566  | 9  | US-09-770-445-478   | Sequence 478, App |
| 30 | 420.5 | 36.6 | 896  | 10 | US-09-764-868-487   | Sequence 487, App |
| 31 | 413.5 | 36.0 | 1127 | 9  | US-09-925-302-340   | Sequence 340, App |
| 32 | 413.5 | 36.0 | 3124 | 10 | US-10-198-846-11497 | Sequence 11497, A |
| 33 | 413   | 35.9 | 2873 | 9  | US-09-770-445-487   | Sequence 487, App |
| 34 | 411   | 35.7 | 894  | 10 | US-09-880-107-3393  | Sequence 3393, Ap |
| 35 | 408   | 35.5 | 1061 | 10 | US-10-037-270-646   | Sequence 646, App |
| 36 | 403.5 | 35.1 | 1069 | 9  | US-09-917-800A-1461 | Sequence 1461, Ap |
| 37 | 401.5 | 34.9 | 585  | 10 | US-09-728-445-652   | Sequence 652, App |
| 38 | 399.5 | 34.7 | 771  | 10 | US-09-938-842A-2113 | Sequence 2113, Ap |
| 39 | 392.5 | 34.1 | 654  | 9  | US-10-198-846-13439 | Sequence 13439, A |
| 40 | 388   | 33.7 | 725  | 9  | US-09-834-975-1     | Sequence 1, Appl  |
| 41 | 386   | 33.6 | 2771 | 10 | US-09-917-800A-1426 | Sequence 1426, Ap |
| 42 | 385   | 33.5 | 857  | 10 | US-09-911-904-150   | Sequence 150, App |
| 43 | 383   | 33.3 | 514  | 9  | US-09-924-035A-794  | Sequence 794, App |
| 44 | 383   | 33.3 | 538  | 10 | US-09-954-456-563   | Sequence 563, App |
| 45 | 383   | 33.3 | 1042 | 10 |                     |                   |

#### ALIGNMENTS

##### RESULT 1

US-10-037-270-959

Sequence 959, Application US/10037270

Publication No. US20030104529A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aildong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: Dmanac, Radoje T.

TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and

FILE REFERENCE: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/10/037,270

CURRENT FILING DATE: 2002-01-04

PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/488,725  
 ; PRIOR FILING DATE: 2000-01-21  
 ; NUMBER OF SEQ ID NOS: 1104  
 ; SOFTWARE: pt-fl\_genes Version 1.0  
 ; SEQ ID NO 959  
 ; LENGTH: 1106  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (124)..(699)  
 ; US-10-037-270-959

Alignment Scores:  
 Pred. No.: 4,31e-139 Length: 1106  
 Score: 1150.00 Matches: 223  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-817-199B-2 (1-223) x US-10-037-270-959 (1-1106)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 DB 28 ATGACGGGCACGCCAGCGCGCTTCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCGC 87  
 QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 DB 88 CCTCGAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTCTCTGGGAGACACAGCGCTC 147  
 QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 DB 148 GGCMAAACATGTTCTCTGATCCAAATCAAGACGGGGCTTCTCTCGGGAACCTTCATA 207  
 QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
 DB 208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGTGATGGTGGCGTGAAGTGAAG 267  
 QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 DB 268 CTGCACATCTGGACACCCCTGGCGAGGAACGGTTCGGAAGCGTCACCCATCTATTATC 327  
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 DB 328 AGAGATGCTCAGCGCTTGTCTGTGTATGATCATCACCACCAAAATCTCTTTTCGACAC 387  
 QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 DB 388 ATCAGGGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 447  
 QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 DB 448 CTAGGCAACAAGCGGATATGAGCAGCAAGAGTATCGGTTCGCAACAGCGAGAGACC 507  
 QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 DB 508 TTGGCCAGGAGTACGGTGTCTCTCTGGAGACAGCGCCCAAGACTGGCATGATGATG 567  
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 DB 568 GAGTTAGCTCTTCTGGCCATCGCAAGGAACCTGAATACCGGGCCGGGATCAGGCGGAT 627  
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 DB 628 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTCCTGCTGC 687  
 QY 221 SerPheMet 223  
 DB 688 TCCTTCATG 696

RESULT 2

US-09-817-199A-1  
 ; Sequence 1, Application US/09817199A  
 ; Patent No. US20020142380A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SHAO, Wei et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
 ; FILE REFERENCE: CLO01187  
 ; CURRENT APPLICATION NUMBER: US/09/817,199A  
 ; CURRENT FILING DATE: 2001-03-27  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2674  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-817-199A-1

Alignment Scores:  
 Pred. No.: 1.6e-138 Length: 2674  
 Score: 1150.00 Matches: 223  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-817-199B-2 (1-223) x US-09-817-199A-1 (1-2674)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 DB 42 ATGACGGGCACGCCAGCGCGCTTCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCGC 101  
 QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 DB 102 CCTCGAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTCTCTGGGAGACACAGCGCTC 161  
 QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 DB 162 GGCMAAACATGTTCTCTGATCCAAATCAAGACGGGGCTTCTCTCGGGAACCTTCATA 221  
 QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
 DB 222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGTGATGGTGGCGTGAAGTGAAG 281  
 QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 DB 282 CTGCACATCTGGACACCCCTGGCGAGGAACGGTTCGGAAGCGTCACCCATCTATTATC 341  
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 DB 342 AGAGATGCTCAGCGCTTGTCTGTGTATGATCATCACCACCAAAATCTCTTTTCGACAC 401  
 QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 DB 402 ATCAGGGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 461  
 QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 DB 462 CTAGGCAACAAGCGGATATGAGCAGCAAGAGTATCGGTTCGCAACAGCGAGAGACC 521  
 QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 DB 522 TTGGCCAGGAGTACGGTGTCTCTCTGGAGACAGCGCCCAAGACTGGCATGATGATG 581  
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 DB 582 GAGTTAGCTCTTCTGGCCATCGCAAGGAACCTGAATACCGGGCCGGGATCAGGCGGAT 641  
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 DB 642 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTCCTGCTGC 701

```
Oy 221 SerPheMet 223
Db 702 TCCTTCATG 710

RESULT 3
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: Human G-Proteins
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)....(699)
US-09-794-257-13

Alignment Scores:
Pred. No.: 1.94e-138 Length: 1116
Score: 1145.00 Matches: 222
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
Gaps: 10

US-09-817-199B-2 (1-223) x US-09-794-257-13 (1-1116)

Oy 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db 31 ACGGGCAGCGCGCGCGTGGCCACCGGATGGCGAGGCGCCCGGAGCGCTCCCGCC 90

Oy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 91 TGCAGTCCGAGCTACGACCTACCGGCAAGGTGATGCTTCTGGAGACACAGGCGTCCGG 150

Oy 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 151 AAAACATGTTCTCTGATCCCAATTCAAGACGGGCGCTTCCTGTCGGACCTTCATAGCC 210

Oy 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 211 ACCGTCCGCGATAGACTTCAGGAACAAGGTGTGACTGTGGATGGCGTGAGAGTGAAGCTG 270

Oy 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 271 CAGATCTGGGACACCGCTGGCGAGAACGGTTCGGAAGCGTTCACCAATGCTTATACAGA 330

Oy 102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGGCGCTTCTGCTGTATGACATCACCAACAAATCTCTTCGACACATC 390

Oy 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db 391 AGGCGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATG 450

Oy 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCAACAGCGGATATGACGAGCAAGAGTATGATGATGATGATGATGATGATGATGATG 510

Oy 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCCAGGAGTACGGTGTTCCTCTCTGGAGACCGCCAGAGCTGGCATGAATGTGGAG 570

Oy 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCCTTTCTGGCCATCGCAAGAACTGAATACCGGGCGGCATCAGCGGATGAG 630

Oy 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
Db 631 CCCAGCTTCAGATCCGAGACTATGTAGATCCCAAGAGCGCTCCAGCTGCTGCTCC 690

Oy 222 PheMet 223
Db 691 TTCATG 696

RESULT 4
US-09-764-868-71
; Sequence 71, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-71

Alignment Scores:
Pred. No.: 9.31e-138 Length: 2623
Score: 1144.00 Matches: 222
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 99.48% Indels: 0
Gaps: 9

US-09-817-199B-2 (1-223) x US-09-764-868-71 (1-2623)

Oy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 12 ATGACGGGACCGCGCGCGTGGCCACCGGATGGCGAGGCGCCCGGAGCGTCCCGC 71

Oy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 72 CCTGCGAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 131

Oy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 132 GGCAAAACATGTTCTCTGATCCCAATTCAAGACGGGCGCTTCCTGTCCGGAACCTTCATA 191

Oy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 192 GCCACCTTCGCGATAGACTTCAGCAACAAGGTGTGACTGTGGATGGCGTGAGAGTGAAG 251

Oy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 252 CTGCAGATCTGGGACACCGCTGGCGAGGAACGGTTCGGAACGGTTCACCATGCTTATATAC 311

Oy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 312 AGAGATGCTCAGGCGCTTCTGCTGTATGACATCACCAACAAATCTCTTTCGACACAC 371

Oy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 372 ATCAGGCGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 431

Oy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 432 CTAGGCAACAAGGCGGATATGAGCAGCGGAAAGAGTGTATCCGTTCGGAACGAGAGACC 491
```

|    |     |  |     |
|----|-----|--|-----|
| Qy | 161 | LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal | 180 |
|    |     |  |     |
| Db | 492 | TTGGCCAGGAGTACGGTGTCCCTTCTGGAGACCAGCGCCAAAGACTGGCATGAATGTG   | 551 |
|    |     |  |     |
| Qy | 181 | GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp | 200 |
|    |     |  |     |
| Db | 552 | GAGTTAGCCTTTCTGGCCATCGCCAAAGAACTGAAATACCGGCCGGGCATCAGCGGAT   | 611 |
|    |     |  |     |
| Qy | 201 | GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys | 220 |
|    |     |  |     |
| Db | 612 | GAGCCAGGTTCCAGATCCGAGACTATGTAGACTCCAGAGAACGCGCTCCAGCTGCTGC   | 671 |
|    |     |  |     |
| Qy | 221 | SerPheMet  | 223 |
|    |     |  |     |
| Db | 672 | TCCTTCATG  | 680 |

## RESULT 5

```

US-10-051-986-10
; Sequence 10, Application US/10051986
; Patent No. US20020146770A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Hillman, Jennifer L.
;           Tang, Y. Tom
;           Lal, Preeti
;           Guegler, Karl J.
;           Corley, Neil C.
;           Patterson, Chandra
;           Batra, Sajeev
;           Baughn, Mariah R.
;
; TITLE OF INVENTION: RAS PROTEINS
;
; NUMBER OF SEQUENCES: 14
;
; CORRESPONDENCE ADDRESSES: 14
;
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC
;
;

```

```

?
?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Diskette
?
? COMPUTER: IBM Compatible
?
? OPERATING SYSTEM: DOS
?
? SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
?
? CURRENT APPLICATION DATA:
?
? APPLICATION NUMBER: US/10/051.986
?
? FILING DATE: 15-Jan-2002
?
? CLASSIFICATION: <Unknown>
?
?
```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/766,551  
FILING DATE: DECEMBER 12, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <unknown>

```

?       1555A: SOURCE:
? INFORMATION FOR SEQ ID NO: 10:
?     SEQUENCE CHARACTERISTICS:
?       LENGTH: 875 base pairs
?       TYPE: nucleic acid
?       STRANDEDNESS: single
?       TOPOLOGY: linear
? IMMEDIATE SOURCE:
?     LIBRARY: UCMCL5T01
?     CLONE: 1528559
?     SEQUENCE DESCRIPTION: SEQ ID NO: 10 :
? US-10-051-986-10

```

|                   |           |
|-------------------|-----------|
| Alignment Scores: |           |
| Pred. No.:        | 6.01e-138 |
| Length:           | 875       |

|   |         |  |     |
|---|---------|--|-----|
| Score:  | 1140.00 | Matches:   | 221 |
| Percent Similarity:                                 | 100.00% | Conservative:  | 0   |
| Best Local Similarity:                              | 100.00% | Mismatches:  | 0   |
| Query Match:  | 99.13%  | Indels:  | 0   |
| DB:   | 12      | Gaps:  | 0   |
| US-09-817-199B-2 (1-223) x US-10-051-986-10 (1-875) |         |  |     |
| Qy  | 3       | GlyThrProGlyAlaValAlaIaThrArgaspGlyGluAlaProGluArgSerProCys  | 22  |
| Db  | 3       | GGCAGCGCAGCGCGGTGGCCACCGGGATGGCAGGCCCCCGAGCGCTCCCGCCCTGC     | 62  |
| Qy  | 23      | SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys | 42  |
| Db  | 63      | AGTCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGCMAA   | 122 |
| Qy  | 43      | ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr | 62  |
| Db  | 123     | ACATGTTTCCTGATCCCAATTCAAGACGGGGCTTCCTGTCGGAACTTCATAGCCACC    | 182 |
| Qy  | 63      | ValGlyIleAspPheArgAsnLysValValThrValaspGlyValArgValLysLeuGln | 82  |
| Db  | 183     | GTCGGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCTGAGAGTGAGCTGCAG     | 242 |
| Qy  | 83      | IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgasp | 102 |
| Db  | 243     | ATCTGGGACACCGTGGCAGGAACGGTTCGGAAGCGTCACCCATGCTTATTACAGAGAT   | 302 |
| Qy  | 103     | AlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg    | 122 |
| Db  | 303     | GCTCAGCGCTTGCTTCTGTATGATCATCAACAATACTCTTCGACAACATCAGG        | 362 |
| Qy  | 123     | AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly | 142 |
| Db  | 363     | GCCTGGCTCATGTAGATTCATGATATGCCAGAGGAGCGTGTGATCATGTGTGATGGC    | 422 |
| Qy  | 143     | AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla | 162 |
| Db  | 423     | AACAAGGCGGATATGAGCAGCGAAGAGTGATCCGTTCCGAAGACGGAGACCTTGGCC    | 482 |
| Qy  | 163     | ArgGluTyrGlyValProPheLeuIuThrSerAlaLysThrGlyMetAsnValGluLeu  | 182 |
| Db  | 483     | AGGAGATACGGTGTTCCCTTCCTGGAGACACGCGCAAGACGCGGATGAATGTGGATTA   | 542 |
| Qy  | 183     | AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro | 202 |
| Db  | 543     | GCCTTTCTGGCCATCGCCCAAGAACTGAATACCGGCGCGGCATCAGCGCGATGAGCCC   | 602 |
| Qy  | 203     | SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSerPhe    | 222 |
| Db  | 603     | AGCTTCAGATCCGAGACTATGTAGATGCCAGAGAAGCGCTCCAGCTGCTGCTCTTC     | 662 |

## RESULT 6

```

US-09-764-868-493
: Sequence 493, Application US/09764868
: Patent No. US20020168711A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT232
: CURRENT APPLICATION NUMBER: US/09/764,868
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1510
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 493
: LENGTH: 1316
: TYPE: DNA
: ORGANISM: Homo sapiens

```



```
;
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1281)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
; LOCATION: (1299)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-764-868-493

Alignment Scores:
Pred. No.:      8.88e-137      Length:      1316
Score:          1133.00      Matches:      220
Percent Similarity: 99.10%      Conservative: 0
Best Local Similarity: 99.10%      Mismatches: 2
Query Match:      98.52%      Indels:      0
DB:              9          Gaps:      0

US-09-817-199b-2 (1-223) x US-09-764-868-493 (1-1316)
QY  2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db  1 ACGGGCACGCCAGCGCCGTTGCCACCCGGGATGGCGAGGCCGCCGAGCGCTCCCGCCC 60
QY  22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db  61 TGCAGTCCGAGCTACGACCTCAGCGGCAGGTGATGCTTCTGGAGACACAGCGGTGGC 120
QY  42 LysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db  121 AAAACATGTTCTGTATCCAAATTCAAAGACGGGCGCTTCTGTCGGGAACCTTCATAGCC 180
QY  62 ThrValGlyLysPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db  181 ACCGTGGCATAGACTTCAGSAACAAGGTGGTACTGTGGATGGCGTGAAGAGTGAAGCTG 240
QY  82 GlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db  241 CAGATCTGGACACCGCTGGCGAGGAGCGGTTCGGAAGCGTCAACCATGCTTATACAGA 300
QY  102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db  301 GATGCTCAGCGCTTCTGCTGTATGACATCACCACCAAAATCTTCTTCGACAAATC 360
QY  122 ArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db  361 AGGCGCTGGCTCACTAGATTCATGAGTATGCCAGAGGACGCTGGTGCATGCTGCTA 420
QY  142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db  421 GGCAACAAGCGGATATGACAGCGGAAGAGTATCGGTTCCGAAGACGAGAGACCTTG 480
QY  162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db  481 GCCAGGAGTACGKGTTCCCTTCTGAGACACCGCCCAAGACTGGCATGTAATGTGGAG 540
QY  182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db  541 TTAGCTTTCTGGCCATCGCAAGAACTGAAATACCGCGCGGCGATCAGCGGGATGAG 600
QY  202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSer 221
Db  601 CCCAGCTCCAGATCCGAGACTATGTAKACTCCAGAAAGCGCTCCAGCTGCTGCTCC 660
QY  222 PheMet 223
Db  661 TTCATG 666

RESULT 7
US-09-794-257-15
; Sequence 15, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel

;
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 576
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-15

Alignment Scores:
Pred. No.:      4.15e-117      Length:      576
Score:          977.00      Matches:      191
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      84.96%      Indels:      0
DB:              10          Gaps:      0

US-09-817-199b-2 (1-223) x US-09-794-257-15 (1-576)
QY  33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52
Db  1 ATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTCTCTGTATCCAAATTCAAAGACGG 60
QY  53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyLysAspPheArgAsnLysValVal 72
Db  61 GCCTTCTGTCGGGAACCTTCATAGCCCGCTGGCATAGACTTCAGGAACAGAGTGGTG 120
QY  73 ThrValAspGlyValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPhe 92
Db  121 ACTGTGGATGCGCTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGACGAGACGCTTC 180
QY  93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db  181 CGAAGCGCTCAACCATGCTTATTACAGAGATGCTCAGCGCTTGTCTGTATGACATC 240
QY  113 ThrAsnLysSerSerPheAspAsnIleArgAlaTyrLeuThrGluIleHisGluTyrAla 132
Db  241 ACCAACAAATCTCTTCGACACATCAGCGCTGGCTGCTGAGATTCATGATGATGCC 300
QY  133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
Db  301 CAGAGGACGCTGGTGCATGCTAGTACCAACAAGCGGATATGACGAGCAAGAGTG 360
QY  153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db  361 ATCCGTTCCGAAGACGAGAGACCTTGGCCAGGAGTACGCTGTCTCTCTCGAGACC 420
QY  173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db  421 AGCCCAAGACTGGCATGATGAGAGTTAGCTTCTGGCCATCGCCAGGAACTGAAA 480
QY  193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db  481 TACCGCGCGGCGATCAGCGGATGAGCCAGCTTCAGAGCTCCAGACTATGTAGAGTCC 540
QY  213 GlnLysLysArgSerSerCysSerPheMet 223
Db  541 CAGAAGAAGCGCTCCAGCTGCTGCTCTTCATG 573

RESULT 8
US-09-867-550-1811
; Sequence 1811, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehrahan, Fuad,
; APPLICANT: Conley, Pamela
```

```
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1811
```

```
Alignment Scores:
Pred. No.: 5,47e-84 Length: 447
Score: 720.00 Matches: 138
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 62.61% Indels: 0
DB: 10 Gaps: 0

US-09-817-199B-2 (1-223) x US-09-867-550-1811 (1-447)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
DB 34 ATGACGGGACGCCAGCGCCGTTGCCACCCGGGATGGCGAGGCCCGCGCTCCCGG 93
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
DB 94 CCTGAGTCCGAGCTACGACTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGGTC 153
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheLe 60
DB 154 GGCANAACATGTTCTGTATCCAAATTCAAAGACGGGCGCTTCTCTCCGGAACCTTCATA 213
QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
DB 214 GCACCGCTGGCATAGACTTCAGGACAAGGTGGTGTGCTGGATGGCTGAGAGTGAG 273
QY 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
DB 274 CTGCAGATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGTCAACCCATGCTTATTAC 333
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
DB 334 AGAGATGCTCAGGCGCTTGTCTGTGTATGATCATCACCACCAAAATCTTCTTCGACAC 393
QY 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIle 138
DB 394 ATCAGGCGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 447
```

RESULT 9

```
US-09-918-995-648
; Sequence 648, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 648
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) (434)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-648
```

```
Alignment Scores:
Pred. No.: 1.07e-60 Length: 434
Score: 540.00 Matches: 99
Percent Similarity: 93.39% Conservative: 14
Best Local Similarity: 81.82% Mismatches: 8
Query Match: 46.96% Indels: 0
DB: 9 Gaps: 0
```

US-09-817-199B-2 (1-223) x US-09-918-995-648 (1-434)

```
QY 26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
DB 67 TACGACGTCGCTTCAAGGTCTGCTGGTGGGACTCGGGTGGGGAAGACCTGTCTG 126
QY 46 LeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyLe 65
DB 127 CTGGTGGATTCAAGGATGGTCTTCTCTGGCGGGACCTTCATCTCCACCGTAGGCATT 186
QY 66 AspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTyrAsp 85
DB 187 GACTTCGGAACAAGTCTCGAGCTGGATGCTGGAAGTGAAGTGCAGATGTGGGAC 246
QY 86 ThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAla 105
DB 247 ACAGTGTGTGAGGAGCGGTTCCGAGTGTATCCCATGCTACTACCGGATGCTCATGCT 306
QY 106 LeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTyrLeu 125
DB 307 CTGCTGCTGCTCTACGATGTCACCAACGAGCCCTCTTGACAACTCCAGGCTGGCTG 366
QY 126 ThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAla 145
DB 367 ACCGAGATCCACGAGTAGCTAGCCACGACGACGCTGCTGCTGCTGCTGCTGCTGCTG 426
QY 146 Asp 146
DB 427 GAC 429
```

RESULT 10

```
US-09-764-868-75
; Sequence 75, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 964
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (806)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (898)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
; NAME/KEY: SITE
; LOCATION: (918)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (924)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (952)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (959)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-75

Alignment Scores:
Pred. No.: 1.56e-59 Length: 964
Score: 535.00 Matches: 98
Percent Similarity: 85.40% Conservative: 19
Best Local Similarity: 71.53% Mismatches: 20
Query Match: 46.52% Indels: 0
DB: 9 Gaps: 0

US-09-817-199B-2 (1-223) x US-09-764-868-75 (1-964)

Qy 84 TTPAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAla 103
Db 3 TGGGACACAGCTGGTCAGGAGCGGTTCGCGAGTGTTTACCCTACTACCGGGATGCT 62
Qy 104 GlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsnIleArgAla 123
Db 63 CATGCTCTGCTGCTCTACGATGTCACCAAGGCCCTCTTTCACACATCCAGGCC 122
Qy 124 TrpLeuThrGlnLeuHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsn 143
Db 123 TGGGTGACCGAGATCCACGAGTACGCCAGCAGCTGGCGCTCATGCTGCTGGGAAC 182
Qy 144 LysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArg 163
Db 183 AGGTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
Qy 164 GluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAla 183
Db 243 GAGTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302
Qy 184 PheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSer 203
Db 303 TTCACAGCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCGCAGCGCGCGC 362
Qy 204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 363 TTCGGCTGCATGATTACGTTAAGAGGGGGTTCGAGGGGGCCCTCCTGCTGC 413

RESULT 11
US-09-834-975-879
; Sequence 879, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Bolt, Andrew
; APPLICANT: Brown, Jeffrey
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 879
; LENGTH: 2497
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)- (2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-879

Alignment Scores:
Pred. No.: 7.09e-56 Length: 2497
Score: 511.50 Matches: 106
Percent Similarity: 62.82% Conservative: 41
Best Local Similarity: 45.30% Mismatches: 68
Query Match: 44.48% Indels: 19
DB: 10 Gaps: 4

US-09-817-199B-2 (1-223) x US-09-834-975-879 (1-2497)

Qy 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
Db 8 GGCACGCC-----CCTCGCGCGCGCGCCCTCCCGCCTCTCTCCACCGCCT 55
Qy 22 -----CysSerProSerTyrAspLeuThr 29
Db 56 CCTCTGCTGCCCGCTCAGAGCGCGGAGGAGATGCGGAGAGCGTACGATTATCTC 115
Qy 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
Db 116 TTCAGCTCTGCTGATCGCGGACTCGGGGTAGCGAGACCTCGCTCTCTCCGCTTC 175
Qy 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
Db 176 TCAGAGGACGCGCTTC---AACACCACCTTCATCTCCACCATCGGAATGATTAAAAAT 232
Qy 70 LysValValThrValAspGlyValArgValLysGlnIleTyrPAspThrAlaGlyGln 89
Db 233 AGAAGTAGAGTAAGTAGAAGAAATTAAGCTTCAGATATGGGACACACGGGTCAG 292
Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
Db 293 GAAAGATTCCGAACAATCAGCAGACGCTACTACAGAGGAGCCATGGGCATTTATGCTG 352
Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTyrPLeuThrGluIleHis 129
Db 353 TATGACATCAAAATGAAATCCTTTGACAATATTAAATTTGATATGATATGAATGAC 412
Qy 130 GluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
Db 413 GAGCATGCTCTTCGATGTCGAAAGATGATCTGGGTAAACAATGTGATATGAATGAC 472
Qy 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
Db 473 AAAAGACAAAGTTCAAAAGAAAGAGGGGAGAGCTAGCAATTTGACTATGGGATTAATTC 532
Qy 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
Db 533 TTGGAGCAACGCGCAAAATCCAGTGCAATATGAAGAGGAGCATTTTACACTCTGCACGA 592
Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
Db 593 GATATATGACAAATCAACAGAAAATGAATGACAGCAATTCACAGGACGAGGTGGA 652
Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
Db 653 CCAGTGAATAAACAAGAAACCGATCAAGAAAGACCGAGTTTC 694

RESULT 12
US-09-834-975-885
; Sequence 885, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
```





## Alignment Scores:

Pred. No.: 5,77e-56 Length: 881  
 Score: 507.00 Matches: 95  
 Percent Similarity: 69.38% Conservative: 50  
 Best Local Similarity: 45.45% Mismatches: 56  
 Query Match: 44.09% Indels: 8  
 DB: 10 Gaps: 3

US-09-817-199B-2 (1-223) x US-09-770-445-529 (1-881)

```

QY 18 ArgSerProProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAsp 37
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 48 AGATCGAAACCAACCAATCCCTGAGTACGACTATCTTTCAAGCTCCTGCTTATCGGGGAT 107
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 38 ThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGly 57
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 108 TCTGGCGTAGGCAAGCTCTGCTCTTTTGAGATTCTCTGATGATTCTTATGTAGAA--- 164
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 58 ThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyVal 77
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 165 AGTTACATTAGCATTATGGAGTCGATTTTAAATTAGGACTGTGGAACAAGATGGCAAA 224
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 78 ArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHis 97
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 225 ACAATTAAGCTCCAAATTTGGACACTGCTGTCAGAACGGTTCAGGACTATTACTAGC 284
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 98 AlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSer 117
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 285 AGTTACTACCTGGGGCACATGGGAATTATTATGTCTACGATGTCACAGATGAAGAAAGC 344
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 118 PheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValVal 137
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 345 TTCAATAATGTCACCAATGCTGAGTGAATTGATCGTTATGCTAGTACAAATGTCAAC 404
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 138 IleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAsp 157
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 405 AAACCTCCTGTGTGAAACAAGCTGATCTTACTGAAAACAGAGCCATTCCTTATGAAACT 464
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 158 GlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGly 177
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 465 GCCAAGGCTTTGCCGATGAATCGGGATTCCTTTATGGAGACTAGTGCAAAAGATGCT 524
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 178 MetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg----- 194
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 525 ACAACGTAGAACAGGCTTTTCATGGCAATGTCATCCATCAAGAGAGATGGCTAGC 584
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 195 -----AlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 585 CAACCAAGCTGGGAATAATGCAAGACCACCGCGTCAGATCAGAGGACAGCCTGTGGCA 644
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 213 GlnLysLysArgSerSerCysCysSer 221
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 645 CAGAG-----AACGGCTGCTGTCA 665
  
```

Search completed: June 23, 2003, 21:35:26

Job time : 163 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 23, 2003, 19:39:18 ; Search time 1382 Seconds  
(without alignments)  
2613.309 Million cell updates/sec

Title: US-09-817-199b-2  
Perfect score: 1150  
Sequence: 1 MTCTPGNAVTRDEAPERSP.....FQIRDYVESOKRSCCSFM 223

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 6.0 , Xgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O/cgn2\_1/USPTO.spool/US09817199/runat\_18062003\_145102\_801/app\_query.fasta\_1.391  
-DB-EST -QFMT-fastap -SUFFIX-p2n.rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0  
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45  
-DOCALIGN-200 -THR SCORE-pct -THR MAX-100 -THR MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTPWT-pto -NORW-ext -HEAPSIZ-500 -MINLEN-0 -MAXLEN-2000000000  
-USER-US09817199 -ECGN\_1\_1906 -runat\_18062003\_145102\_801 -NCPU-6 -ICPU-3  
-NO\_MMMap -LARGEQUERY -NEG\_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG  
-DEV\_TIMEOUT-120 -WARN\_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOPOP-6  
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_oth:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID       | Description        |
|------------|-------|---------------|--------|----|----------|--------------------|
| 1          | 1150  | 100.0         | 1014   | 14 | BM921365 | BM921365 AGENCOURT |
| 2          | 1137  | 98.9          | 1051   | 12 | RG283602 | RG283602 602407930 |
| 3          | 1134  | 98.6          | 887    | 9  | AL522282 | AL522282 AL522282  |
| 4          | 1058  | 92.0          | 676    | 10 | BB598938 | BB598938 BB598938  |
| 5          | 1020  | 88.7          | 651    | 10 | BB633978 | BB633978 BB633978  |
| 6          | 1006  | 87.5          | 740    | 13 | BI767046 | BI767046 603054267 |
| 7          | 1001  | 87.0          | 655    | 10 | BB635649 | BB635649 BB635649  |
| 8          | 856.5 | 74.5          | 749    | 9  | AL559085 | AL559085 AL559085  |
| 9          | 856   | 74.4          | 733    | 10 | BB619625 | BB619625 BB619625  |
| 10         | 852   | 74.1          | 522    | 13 | BM151643 | BM151643 TCBAPE10  |
| 11         | 829   | 72.1          | 736    | 13 | EG934396 | EG934396 SKI-0698  |
| 12         | 825   | 71.7          | 676    | 10 | BB206788 | BB206788 BB206788  |
| 13         | 790   | 68.7          | 553    | 12 | BE755280 | BE755280 209068 MA |
| 14         | 763   | 66.3          | 488    | 13 | BM149118 | BM149118 TCAAP2E63 |
| 15         | 760   | 66.1          | 888    | 14 | EQ687186 | EQ687186 AGENCOURT |
| 16         | 760   | 66.1          | 953    | 12 | BE906081 | BE906081 601497118 |
| 17         | 754   | 65.6          | 895    | 14 | EQ688147 | EQ688147 AGENCOURT |
| 18         | 752   | 65.4          | 579    | 10 | BB621927 | BB621927 BB621927  |
| 19         | 740   | 64.3          | 657    | 10 | BB196489 | BB196489 BB196489  |
| 20         | 739   | 64.3          | 455    | 10 | AW464449 | AW464449 BP230015B |
| 21         | 728   | 63.3          | 691    | 12 | BG253976 | BG253976 602366910 |
| 22         | 721.5 | 62.7          | 672    | 13 | BM432040 | BM432040 P922n.pko |
| 23         | 650   | 56.5          | 701    | 10 | AW914157 | AW914157 EST345461 |
| 24         | 646   | 56.2          | 708    | 9  | AJ442405 | AJ442405 AJ442405  |
| 25         | 622   | 54.1          | 780    | 12 | EG784310 | EG784310 SEAU0C004 |
| 26         | 610   | 53.7          | 443    | 10 | AW956878 | AW956878 EST368948 |
| 27         | 610   | 53.0          | 380    | 10 | BE241772 | BE241772 TCAAP2E02 |
| 28         | 607   | 52.8          | 386    | 10 | AW484587 | AW484587 61493 MAR |
| 29         | 588   | 51.1          | 820    | 13 | BI946873 | BI946873 603177832 |
| 30         | 582   | 50.6          | 782    | 13 | BI545231 | BI545231 603187353 |
| 31         | 581.5 | 50.6          | 658    | 10 | AV697217 | AV697217 AV697217  |
| 32         | 581   | 50.5          | 368    | 13 | BM152052 | BM152052 TCBAPE12  |
| 33         | 580.5 | 50.5          | 648    | 10 | AV697216 | AV697216 AV697216  |
| 34         | 580.5 | 50.5          | 684    | 10 | AV692701 | AV692701 AV692701  |
| 35         | 580.5 | 50.5          | 697    | 10 | AV692689 | AV692689 AV692689  |
| 36         | 576.5 | 50.1          | 678    | 10 | AV685389 | AV685389 AV685389  |
| 37         | 572   | 49.7          | 668    | 10 | AV688197 | AV688197 AV688197  |
| 38         | 557.5 | 48.5          | 647    | 10 | AV698293 | AV698293 AV698293  |
| 39         | 557   | 48.4          | 678    | 10 | AV686507 | AV686507 AV686507  |
| 40         | 548   | 47.7          | 692    | 12 | EG699266 | EG699266 602678938 |
| 41         | 535   | 46.5          | 866    | 9  | AU079251 | AU079251 AU079251  |
| 42         | 530   | 46.1          | 592    | 13 | BM023074 | BM023074 1d65h09.y |
| 43         | 528.5 | 46.0          | 767    | 12 | BF608560 | BF608560 MY1 00150 |
| 44         | 528.5 | 46.0          | 939    | 14 | BQ947443 | BQ947443 AGENCOURT |
| 45         | 527   | 45.8          | 404    | 13 | BM149006 | BM149006 TCAAP2E57 |

ALIGNMENTS

RESULT 1  
BM921365  
LOCUS BM921365  
DEFINITION AGENCOURT\_6626159 NIH\_MGC\_115 Homo sapiens CDNA clone IMAGE:5752779  
5', mRNA sequence.  
ACCESSION BM921365  
VERSION BM921365.1 GI:19371744  
KEYWORDS EST.  
SOURCE human  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1014)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12787 row: 1 column: 04  
High quality sequence stop: 735.  
Location/Qualifiers

## FEATURES

source

1. .1014  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5752779"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"

/note="organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

BASE COUNT 237 a 288 c 295 g 194 t

## ORIGIN

## Alignment Scores:

Pred. No.: 6.16e-139 Length: 1014  
Score: 1150.00 Matches: 223  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-817-199b-2 (1-223) x BM921365 (1-1014)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
|||||  
DB 48 ATGACGGGACCCAGGCGCCCTTCCACCCGGGATGGGAGGCCCGGCGCTCCCG 107  
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
|||||  
DB 108 CCCTGCAGTCCGAGCTACGACCTCAGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 167  
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheLeu 60  
|||||  
DB 168 GCGAACAATGTTCTGTATCCATTAAGACGGGCTTCTGTCCGAACTTCATA 227  
QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
|||||  
DB 228 GCCACGTCGGCATAGACTTCAGAACAGGTGGTGAAGTGTGATGGGTGAGAGTGAAG 287  
QY 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
|||||  
DB 288 CTGCAGATCTGGGACACCCCTGGGAGGAACGGTTCCGAAGCGTCAACCATCTATTATAC 347  
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
|||||  
DB 348 AGAGATGCTCAGCGCTTGTCTGTGTATGATACATCACCACAAATCTCTTTGCACAAAC 407  
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
|||||  
DB 408 ATCAGGCTGCTCACTCAGATTCATGATGATGCCAGAGGACGTGGTGTATGATGCTG 467  
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
|||||  
DB 468 CTAGGCAACAAGCGGATATGACGCGAAGAGATGATCCGTTCGGAAGACGAGAGACC 527

QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
|||||  
DB 528 TTGCCAGGAGTAGTACGGTGTCTCTCTGAGACCGCCAGAGCTGGCATGAATGTG 587  
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
|||||  
DB 588 GAGTAGCTTCTTGGCCATCCGCAAGGAATGAATACCGGGCGGCATCAGGCGGAT 647  
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
|||||  
DB 648 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCTGCTGC 707  
QY 221 SerPheMet 223  
|||||  
DB 708 TCCTTCATG 716

## RESULT 2

LOCUS BG283602 1051 bp mRNA linear EST 21-FEB-2001  
DEFINITION 602407930F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4520191 5',  
mRNA sequence.  
ACCESSION BG283602  
VERSION BG283602.1 GI:13033710  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1051)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DRP

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10417 row: k column: 08  
High quality sequence stop: 782.  
Location/Qualifiers

## FEATURES

source

1. .1051  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4520191"  
/clone\_lib="NIH\_MGC\_91"  
/tissue\_type="adenocarcinoma, cell\_line"  
/lab\_host="DH10B (phage-resistant)"  
/note="organ: prostate; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.4 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 275 a 287 c 314 g 175 t

## ORIGIN

## Alignment Scores:

Pred. No.: 3.24e-137 Length: 1051  
Score: 1137.00 Matches: 223  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 0  
Query Match: 98.87% Indels: 1  
DB: 12 Gaps: 0

US-09-817-199b-2 (1-223) x BG283602 (1-1051)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
|||||  
DB 3 ATGACGGGACCCAGGCGCCCTTCCACCCGGGATGGGAGGCCCGGCGCTCCCG 62



```

QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
DB 63 CCCTGCAGTCGAGCTACGACCTCAGCGGCAAGGTATGCTTCTGGGAGACACAGGCGTC 122
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheile 60
DB 123 GGCANACATGTTCTTCATCCAAATTCNAAGACGGGCCCTTCCTGTCGGAACCTTCATA 182
QY 61 AlaThrValGlyLysPheAspPheAsnLysValThrValAspGlyValArgValLys 80
DB 183 GCCACCGTCGCATAGACTTCAGGAACAAGTGTGACTGTGGATGCGTGAGAGTGAAG 242
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
DB 243 CTCAGATCTGGGACACCGCTGGCGAGGAACGGTTCCGAAGCGCTACCCATGCTTATATAC 302
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
DB 303 AGAGATGCTCAGCGCTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTCGACAC 362
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
DB 363 ATCAGGCGCTGGCTCACTAGATTCATGAGTATGCCAGAGGACGTGTGATCATGCTG 422
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
DB 423 CTAGGCACAAAGCGGATATGAGCAGCGAAAGAGTATCCGTTCCGAAGACGGAGAC 482
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
DB 483 TTGGCCAGGAGTACGGTGTCCCTTCTGGAGACCGCCAGACCTGTCATGAATGTG 542
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
DB 543 GAGTTAGCCCTTTCTGGCCATCGCAAGGAACCTGAATACCGGGCGGCATCAGCGGAT 602
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGln-LysLysArgSerSerCysCy 220
DB 603 GAGCCCAAGCTTCCAGATCCAGACTATGTAGTCCAGAAAGAGTCCAGAAAGAGCGCTCCAGCTGCTG 662
QY 220 sSerPheMet 223
DB 663 CTCCTTCATG 672

```

## RESULT 3

```

AL522282
LOCUS AL522282 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 5
DEFINITION AL522282 887 bp mRNA linear EST 13-FEB-2001
prime_mRNA sequence.

```

```

ACCESSION AL522282
VERSION AL522282.1 GI:12785775

```

```

KEYWORDS EST.
SOURCE human.

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 887)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

```

```

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

## FEATURES

## source

```

1..887
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB008YF08"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"

```

## /lab\_host="DH10B"

```

/notes=Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

```

BASE COUNT 203 a 256 c 257 g 168 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 6,08e-137 Length: 887
Score: 1134.00 Matches: 220
Percent Similarity: 99.10% Conservative: 1
Best Local Similarity: 98.65% Mismatches: 2
Query Match: 98.61% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-817-199B-2 (1-223) x AL522282 (1-887)

```

```

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
DB 28 ATGACGGGCGACCGCAGCGCGCTTTCACCCCGGATGGCGAGGCCCGCGAGCGCTCCCG 87
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
DB 88 CCTTGCAGTCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheile 60
DB 148 GGCAAAACATGTTCTTCATCCAAATTCAAAGACGGGCGCTTCTTCGGAACCTTCATA 207
QY 61 AlaThrValGlyLysPheAspPheAsnLysValThrValAspGlyValArgValLys 80
DB 208 GCCACCGTCGCATAGACTTCAGGAACAAGGTGTGACTGTGGATGCGCTGAAAGTGAAG 267
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
DB 268 CTGCAGATCTGGGACACCGCTGGCGAGAACGGTTCGAGCGCTCACCATGCTTATATAC 327
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
DB 328 AGAATGCTCAGGCGCTTCTTCGCTGTATGACATCACCAACAAATCTTCTTCGACAAC 387
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
DB 388 ATCAGGCGCTGGCTCACTAGATTCATGACTATCCCGAGGAGCGCTGGTGTATCATGCTG 447
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
DB 448 CTAGGCACAAAGCGGATATGAGCAGCGCAAAAGTGTATCCGTTCCGAGACGGAGAGACC 507
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
DB 508 TTGGCCAGGAGTACGGTGTTCCTTCTGAGACCGACGCGCAAGACTGGCATGAATCTG 567
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
DB 568 GAGTTAGCCCTTTCTGGCCATCGCCCAAGGAACCTGAATACCGGGCGGCGCATCAGCGGAT 627
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
DB 628 GAGCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAAAGAGCGCTCCAGCTGCTGC 687
QY 221 SerPheMet 223
DB 688 TCCCTTCATG 696

```

## RESULT 4

BB598938

LOCUS

DEFINITION

BB598938 RIKEN full-length enriched, adult pancreas islet cells Mus

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB598938 676 bp mRNA linear EST 26-OCT-2001  
 BB598938 RIKEN full-length enriched, adult pancreas islet cells Mus  
 musculus cDNA clone C820003E14 5', mRNA sequence.  
 BB598938  
 EST.  
 house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 676)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda  
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sano, K., Sano, H., Sasaki  
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)  
 On Dec 1, 2000 this sequence version replaced gi:11507539.  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for  
 further details.  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Location/Qualifiers  
 1. 676  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="C820003E14"  
 /clone\_lib="RIKEN full-length enriched, adult pancreas  
 islet cells"  
 /tissue\_type="pancreas"  
 /cell\_type="islet cells"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: Sali; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia

FEATURES

source

Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5',  
 GAGAGAGAGCGCGCCGCACTCGAGTTTTTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5',  
 GAGAGAGAGATCTCGAGTTTAAATATATCCCCCCCCCC 3']. cDNA  
 was cleaved with BamHI and XhoI. Vector: a modified  
 pBluescript KS(+) after bulk excision from Lambda FIC I.  
 Cells were provided by Tomohiro Kono, Department of Animal  
 Science, Tokyo University of Agriculture, Atsugi City,  
 Kanagawa Prefecture, Japan, whose assistance we gratefully  
 acknowledge."

BASE COUNT 157 a 178 c 198 g 142 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,24e-127 Length: 676  
 Score: 1058.00 Matches: 205  
 Percent Similarity: 95.45% Conservative: 5  
 Best Local Similarity: 93.18% Mismatches: 10  
 Query Match: 92.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-817-199b-2 (1-223) x BB598938 (1-676)

QY 1 MethrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 |||||  
 Db 16 ATGACTGGCACACCCAGGAGCTGCTACCGCTGGGATGGCAGGCCCTCAGCGCTCCCG 75  
 |||||  
 QY 21 ProCysSerProSerAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 |||||  
 Db 76 CCTTCACCGCCGAACTAGTACTCACCAGGAGGTGCTCTCTGGAGACTCGGGCGTC 135  
 |||||  
 QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 |||||  
 Db 136 GGCAAAACCTGTTCTCTGATCCAAATCAAGACGGGGCTTCCTGTCCGGAACCTTCATA 195  
 |||||  
 QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValValLys 80  
 |||||  
 Db 196 GCCACCGTCGGCATAGACTCCAGGAATAAGTGTGACAGTGGTGTCCAGGCGTGAAG 255  
 |||||  
 QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 |||||  
 Db 256 CTTGAGATCTGGGACACTGTCAGGACAGGAGCGCTCCGAGTGTGACCATGCTATTATAC 315  
 |||||  
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 |||||  
 Db 316 CGAGATGCTCAGGCTTTGCTCTGTTGATGATCACCACCAACAGTCTCTCTTTTGACAAC 375  
 |||||  
 QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 |||||  
 Db 376 ATCAGGCGCTGGCTCAGAGAGATTTCATGAGTATCCAGAGAGAGACGTGCTATTATGCT 435  
 |||||  
 QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 |||||  
 Db 436 CTAGGCAACAGCGCGATGTAGCAGCGAAGGGTGATCTCTTGAAGATGGAGAGACA 495  
 |||||  
 QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 |||||  
 Db 496 CTGCGCAGGGAATATGTTGTTCTTCATGGAGACAGTGCACAGAGCTGGCATCAACGTG 555  
 |||||  
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 |||||  
 Db 556 GAGTTGGCTTTCTGGNCATTTGCCAAGGAATACCGTGCAGGAGGAGCGACCTGAT 615  
 |||||  
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 |||||  
 Db 616 GAGCCAGCTTCCAGATCCGAGACTGTGGAGTCCCGAAGAGAGCGCTCCAGCTGTGTC 675

| BASE COUNT   | 151 a     | 168 c  | 192 g | 140 t |
|--|-----------|--|-------|-------|
| ORIGIN   |           |  |       |       |
| Alignment Scores:  |           |  |       |       |
| Pred. No.:   | 2,76e-122 | Length:  | 651   |       |
| Score:   | 1020.00   | Matches:   | 198   |       |
| Percent Similarity:  | 95.75%    | Conservative:  | 5     |       |
| Best Local Similarity:   | 93.40%    | Mismatches:  | 9     |       |
| Query Match:   | 88.70%    | Indels:  | 0     |       |
| DB:  | 10        | Gaps:  | 0     |       |
| US-09-817-199B-2 (1-223) x BB633978 (1-651)                      |           |  |       |       |
| QY   | 1         | MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro | 20    |       |
| Db   | 16        | ATGACTGGCACACACAGGAGCTGTACCGCTGGGGATGGGAGGCCCTGAGCGCTCCCG    | 75    |       |
| QY   | 21        | ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal | 40    |       |
| Db   | 76        | CCCTTCAGCCGGAACACACATCTACCGGCAAGGTGATGCTCTTGAGACTCGGGCGTC    | 135   |       |
| QY   | 41        | GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle | 60    |       |
| Db   | 136       | GGCAAAACCTGTTTCTGATCCAAATCAAAGACGGGCGCTTCCTGTCGGAACCTTCATA   | 195   |       |
| QY   | 61        | AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys | 80    |       |
| Db   | 196       | GCCACCGCTGGCATGACATTCAGGAATAAGTGGTGACAGTGGATGGTCCAGGGGTGAAG  | 255   |       |
| QY   | 81        | LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr | 100   |       |
| Db   | 256       | CTTCACATCTGGACACTGCAGGACAGGAGCGCTTCGCGAGTGTGACCCATGCTTATTAC  | 315   |       |
| QY   | 101       | ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn | 120   |       |
| Db   | 316       | CGAGATGCTCAGGCTTTGCTCTGTGTATGACATCACCAACCACTCTCTTTTGACAAC    | 375   |       |
| QY   | 121       | IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu    | 140   |       |
| Db   | 376       | ATCAGGGCTGGCTCACAGATTCATGATGATGATGCCAGAGACGCTGTGATATGCTT     | 435   |       |
| QY   | 141       | LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr | 160   |       |
| Db   | 436       | CTA5GCAACAAGCCCATGTAAACACAGGAAGGCTGATCCGTTCTGNAGATGAGAGACA   | 495   |       |
| QY   | 161       | LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal | 180   |       |
| Db   | 496       | CTGCCAGGGAATATGCTGTCTCTTCATGGAGACCATGCCCAGACGTCGCATGAACGCTG  | 555   |       |
| QY   | 181       | GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp | 200   |       |
| Db   | 556       | GAGTTGGCTCTTCTGGGCATTCGCAAGGAACATGAATACCGTGCAGGAGGCACCTGAT   | 615   |       |
| QY   | 201       | GluProSerPheGlnIleArgAspTyrValGluSer                         | 212   |       |
| Db   | 616       | GACCCAGCTTTCAGATCCGAGACTATGTCGAGTCC                          | 651   |       |
| RESULT 6   |           |  |       |       |
| BT676046   |           |  |       |       |
| LOCUS  |           |  |       |       |
| DEFINITION   |           |  |       |       |
| 603054267F1 NIH MGC 122 Homo sapiens cDNA clone IMAGE:5203600 5' |           |  |       |       |
| linear   |           |  |       |       |
| EST 25-SEP-2001  |           |  |       |       |
| IMAGE:5203600 5'   |           |  |       |       |

|          |            |             |             |              |            |                  |
|----------|------------|-------------|-------------|--------------|------------|------------------|
| BI767046 | LOCUS      | BI767046    | 740 bp      | mRNA         | linear     | EST 25-SEP-2001  |
|          | DEFINITION | 603054267f1 | NIH MGC 122 | Homo sapiens | cdna clone | IMAGE:5203600 5' |

|            |  |  |     |
|------------|--|--|-----|
| QY         | 100  | rArgAspAlaGlnAlaLeuLeuLeuTyRAspIleThrAsnLysSerSerPheAspAs    | 120 |
| Db         | 301  | CAGAGATGCTCAGGCGCTTCTGCTGTATGACATCACCACAAATCTCTTTTGACAA      | 360 |
| QY         | 120  | nIleArgAlaTrpLeuThrGluIleHisGluTyRAlaGlnArgAspValIleMetL     | 140 |
| Db         | 361  | CATCAGGGCTGCTCCTCAGAGATTCTATGAGTATGCCAGAGGAGGTGTTGATCATGC    | 420 |
| QY         | 140  | eLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyClut  | 160 |
| Db         | 421  | TGCTAGGCACAAAGCGGATATGAGCAGCGCAAGAGTGTATCCGTTCCGAGACGGAGAGA  | 480 |
| QY         | 160  | hIleuAlaArgGluTyRgLyValproPheLeu-GluThrSerAlaLysThr-glyMetAs | 179 |
| Db         | 481  | CTTTGGGAGGTAGTACGGTGTTCCTCTCGGAGAAACAGTCCCAAGACTTGGCATGAA    | 540 |
| QY         | 179  | nValGluLeu-AlaPheLeuAlaIleAlaLysGluLeuLysTyRArgAlaGlyHisGlnA | 199 |
| Db         | 541  | TGTGGAGTTAGGCTTTTCTTGGCATCGCCAGGGAACCTTGAATACCGGCCGGCATCAGG  | 600 |
| QY         | 199  | laAspGluProSerPheGlnIle-ArgAspTyRValGlu-SerGlnLysLysArgSerSe | 218 |
| Db         | 601  | CGATGAACCCAGCTTCCAGATTCCGAGACTGTAGAAGTCCCAAGAAAGCGCTCAAG     | 660 |
| QY         | 218  | r-CysCysSerPheMet 223  |     |
| Db         | 661  | CTTGCTGGTCTTCATG 677   |     |
| RESULT 7   |  |  |     |
| LOCUS      | BB635649   |  |     |
| DEFINITION | BB635649 RIKEN full-length enriched, 0 day neonate thymus Mus        | 655 bp mRNA linear EST 26-OCT-2001                           |     |
| ACCESSION  | BB635649   |  |     |
| VERSION    | BB635649.1   | GI:16471692  |     |
| KEYWORDS   | EST.   |  |     |
| SOURCE     | house mouse  |  |     |
| ORGANISM   | Mus musculus   |  |     |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |  |     |
| AUTHORS    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.   |  |     |
|            | 1 (bases 1 to 655)   |  |     |
|            | Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., |  |     |
|            | Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda    |  |     |
|            | ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,         |  |     |
|            | Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki  |  |     |
|            | ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,     |  |     |
|            | Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,   |  |     |
|            | Muramatsu,M. and Hayashizaki,Y.                                      |  |     |
|            | RIKEN Mouse ESTs (Arakawa,T., et al. 2001)                           |  |     |
|            | Unpublished (2001)   |  |     |
| TITLE      | Contact: Yoshihide Hayashizaki                                       |  |     |
| JOURNAL    | Laboratory for Genome Exploration Research Group, RIKEN Genomic      |  |     |
| COMMENT    | Sciences Center(GSC), Yokohama Institute                             |  |     |
|            | The Institute of Physical and Chemical Research (RIKEN)              |  |     |
|            | 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan    |  |     |
|            | Tel: 81-45-503-9222  |  |     |
|            | Fax: 81-45-503-9216  |  |     |
|            | Email: genome-res@sc.riken.go.jp,                                    |  |     |
|            | URL:http://genome.gsc.riken.go.jp/                                   |  |     |
|            | Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh   |  |     |
|            | ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.           |  |     |
|            | Normalization and subtraction of cap-trapper-selected cDNAs to       |  |     |
|            | prepare full-length cDNA libraries for rapid discovery of new        |  |     |
|            | genes. Genome Res. 10 (10), 1617-1630 (2000)                         |  |     |
|            | waghi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,      |  |     |
|            | Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura   |  |     |
|            | ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and       |  |     |
|            | Hayashizaki,Y.   |  |     |
|            | RIKEN integrated sequence analysis (RISA) system--384-format         |  |     |
|            | sequencing pipeline with 384 multicapillary sequencer. Genome Res.   |  |     |
|            | 10 (11), 1757-1771 (2000)  |  |     |
|            | Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara   |  |     |
|            | ,Y. and Hayashizaki,Y.   |  |     |





```
Db 3 ACTGGCACACTAGGAGCTGTACCGCTGGGATGGCGAGGCCCTGAGCGCTCCCCGCC 62
Qy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 63 TTCAGCCGACACTACGATCTCACCGCAAGGTGATGCTTCTGGAGACTGGGGCTGGC 122
Qy 42 LysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheLeuAla 61
Db 123 AAACCTGTCTCTGATCCCAATTCAAAGACGGGCGCTCTCTCCGGAACCTTCATAGCC 182
Qy 62 ThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 183 ACCGTCGGCATAGACTTCAGGATAAAGTGGTGACAGTGGTGGCCAGGGTGAAGCTT 242
Qy 82 GlnLeuTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 243 CAGATCTGGGACACTGCAGACAGGAGCGCTTCGGCAGTGTGACCCATGCTTATTACCGA 302
Qy 102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspLeuThrAsnLysSerSerPheAspAsnIle 121
Db 303 GATGCTCAAGCTTGTCTCTGTTGATGACATCACCACCAATCTCTTTTGGACAACATA 362
Qy 122 ArgAlaTrpLeuThr-GlnLeuHisGluTyrAlaGlnArgAspValValIleMetLeuLe 141
Db 363 AGGCGCTGGCTTCAAAAGATTCATGAAGTTGCCACAGAGACGTGGTGATATGCTTTT 422
Qy 141 uGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLe 161
Db 423 AGGCAC-AGGCGCGATGTAGACGACGGAAGGGTATCCGTTCTGAAGAAGAAAGACACT 481
Qy 161 uAlaArgGluTyrGlyValProPheLeuGlu-ThrSerAlaLysThrGlyMetAsnValG 181
Db 482 GGCAGGAGAAATAGTGGTGCTCTTTCATGGAGAACCGTGGCCAGACGTGGTGAACCTGG 541
Qy 181 luLeuAla-PheLeuAlaIleAlaLysGlu-LeuLysTyrArgAla-GlyHisGlnAlaA 200
Db 542 AGTTGGCCCTTCTGGCAATTCGCAAGGAATTAATAACCGTCAAGAGGAGCGAGCTCTG 601
Qy 200 spGluProSerPheGlnIleArgAspTyrValGlu-SerGlnLysLysArgSerSerCys 219
Db 602 ATGAACCCCACTTTCAAACCGAGATATGTGGAAGTCCCAAAAAACCGCTCCAACTGG 661
Qy 220 Cys 220
Db 662 TGC 664

RESULT 10
BM151643
LOCUS
DEFINITION
TCBAP10302 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1030, mRNA
sequence.
ACCESSION
BM151643
VERSION
BM151643.1 GI:17175247
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 522)
WEL,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
Günaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
Unpublished (2001)
CONTACT: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Seq primer: M13 primer.
FEATURES
Location/Qualifiers
```

```
source
1. 522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TCBAP1030"
/leukemia_baylor_HGSC_project="TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/notes="Vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCCGAGGAG(T)VN
3'; V-A,C,G; N-A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGCTCGATCCGCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda PSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)."
BASE COUNT 119 a 146 c 158 g 99 t
ORIGIN

Alignment Scores:
Pred No.: 1,6e-100 Length: 522
Score: 852.00 Matches: 165
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.09% Indels: 0
DB: 13 Gaps: 0

US-09-817-199B-2 (1-223) x BM151643 (1-522)
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 19 ATGACGGCAGCCAGCGCGCTTGCACCGCGGATGGCGAGCGCCCGCGCTCCCG 78
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 79 CCCTGCGCTCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGTC 138
Qy 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheTle 60
Db 139 GGCAAAACATGTTCTTCATGATCCAAATTCAGACCGCGGCTTCTCTCCGGAACCTTCATA 198
Qy 61 AlaThrValGlyLysAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 199 GCCACCGTCGCGATAGACTTCAGCAACAAGGTGCTGACTGTGGATGGCGTCAGAGTGAAG 258
Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 259 CTGCAGATCTGGGACACCGCTGGCGAGAACGGTTCGAGAGCGTCACCCATGCTTATTAC 318
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspThrAsnLysSerSerPheAspAsn 120
Db 319 AGACATGCTCAGGCGCTTCTCTGCTGATGACATCACCACCAATCTCTTTCGACAAAC 378
Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 379 ATCAGGCGCTGGCTCACTGAGATTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTG 438
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 439 CTAGGCAACAGCGCGGATATGAGCAGGGAAGAGTGTCCGTTCCGAAGACGAGAGACC 498
Qy 161 LeuAlaArgGluTyr 165
|||||
```



```

Db 499 TTGGCCAGGAGTAC 513
RESULT 11
BC934396/c
LOCUS
DEFINITION
  736 bp mRNA linear EST 02-OCT-2001
  SKI-0698 Atlantic Salmon kidney Salar cDNA clone SKI-0698
  similar to GTPase Rab37, mRNA sequence.
ACCESSION
  BG934396
VERSION
  BG934396.1 GI:15844224
KEYWORDS
  EST.
SOURCE
  Atlantic salmon.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
  1 (bases 1 to 736)
  Douglas, S.E., Tsol, S.C.M., Penny, S., Melville, K., Reith, M.E. and
  Ewart, K.V.
  Expressed Sequence Tags - A Snapshot of the Fish Genome
  Unpublished (2001)
  Contact: Douglas S
  Genome Sciences
  NRC Institute for Marine Biosciences
  1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada
  Tel: (902) 426-4991
  Fax: (902) 426-9413
  Email: susan.douglas@nrc.ca
  Seq primer: SK
  Location/Qualifiers
    1..736
      /organism="Salmo salar"
      /db_xref="taxon:8030"
      /clone="SKI-0698"
      /clone_lib="Atlantic Salmon kidney"
      /env_stages="adult"
      /note="Organ: kidney; Vector: Lambda ZAP II"
BASE COUNT 146 a 220 c 180 g 190 t
ORIGIN

Alignment Scores:
Pred. No.: 2,786-97 Length: 736
Score: 829.00 Matches: 155
Percent Similarity: 90.26% Conservative: 21
Best Local Similarity: 79.49% Mismatches: 19
Query Match: 72.09% Indels: 0
DB: 13 Gaps: 0

US-09-817-199b-2 (1-223) x BG934396 (1-736)

QY 26 TyrAspLeuThrGlyValMetLeuLeuGlyAspThrGlyValGlyLeuThrCysPhe 45
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 666 TTTGATATTGCTTCAAGGTGATGCTGCTGGAGACTCTGCGGTGGGAACACATGTGTG 607
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 46 LeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyLeu 65
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 606 CTGGTGGCTTTAAGACGGGGCTTTTGGGAGGCACACTTCATAGCCACCGCTGGGATA 547
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 66 AspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAsp 85
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 546 GACITTAGATTAAGTTGTGACGGTGGACACACCGAGGTCAACTCCAGATCTGGAT 487
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 86 ThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrrArgAspAlaGlnAla 105
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 486 ACTGAGGACAGGAGAGATTCGCCAGTGTCCAGCAGCGCTACTACAGAGACGACAGGCC 427
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 106 LeuLeuLeuLeuTyrrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeu 125
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 426 CTCCTCTGCTCTATGATCATCCAGCAAGTCACTCTTGGACACATCAGGCGCTGGCTG 367
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 126 ThrGluIleHisGluTyrrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAla 145
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 366 ACTGAATCCATGATGTATGCTCAGAGAGATGTGTGTCATCATGTGTGCTCGGCAACAGACG 307
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

QY 145 AspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyr 165
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 306 GACATGGCCGAGAGAGGGTCAAGCGCTGAGGAGGAGAGCTGCCAAGGATAT 247
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 166 GlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeu 185
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 246 GGAGTCCCTTCATGGAGACCACTGCCAAGACTGGAGTCAAGTCAAGTGGCTTCCCTG 187
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 186 AlaIleAlaLysGluLeuLysTyrrGAlaGlyHisGlnAlaAspGluProSerPheGln 205
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 186 GCTGTAGCAAGAGGAGTGTGAGCAGACAGCTGCCAGCAGCCCAAGGATTTCCAG 127
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 206 IleArgAspTyrrValGluSerGlnLysLysArgSerSerCysCys 220
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 126 ATCCACGACTACATCGCTCTGAGAAGCAGCAAGTCTGGCTGTGT 82
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

## RESULT 12

BB206788

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB206788  
 musculus  
 BB206788.2  
 EST.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 676)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,  
 Muramatsu, M. and Hayashizaki, Y., et al. 2001  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)  
 On Jun 30, 2000 this sequence version replaced gi:8871741.  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alzawa,  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.  
 cDNA library was prepared and sequenced in Mouse Genome





Db 131 CCTGGGGCCGAGCCAGCATCTCAGGGCAAGGTGATGCTTCTCGGAGACTCGGGCGTC 190

QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
|||||  
Db 191 GCAAAACCTGTTCTCTCATCGATCAGAGACGGGGCTTCTCTCGGGACCTTCATA 250

QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
|||||  
Db 251 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGACCGTGGATGGTGTGAGAGTGAA 310

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTrpTyr 100  
|||||  
Db 311 CTGCAGATCTGGACACGCGCAGGCGAGCGGTTCGCGAGTGTACCCATGCTTATATAC 370

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsn 120  
|||||  
Db 371 CGAGAGCGCCAGCGCTTCTCTGTGTACGACATCAGCAACAAATCTTCTTTGACAAAC 430

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140  
|||||  
Db 431 ATCCGGGCTTGCTCCTCAGTATTCAGGATCAGCGCCAGAGGGATGTGTGATCATGCTT 490

QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
|||||  
Db 491 CTGGCAACAAGCGCAGATGTGACACTGAAGGGTGTATCCGCTCAGAGACGAGAGATG 550

QY 161 Leu 161  
|||  
Db 551 CTG 553

RESULT 14  
BM149118  
LOCUS  
DEFINITION  
TCAAP26371 Pediatric acute myelogenous leukemia cell (FAB M1)  
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP6371, mRNA  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
BM149118  
BM149118.1 GI:17170448  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 488)  
Wei, Y., Teang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,  
Guaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.  
Pediatric Leukemia cDNA Sequencing Project (2001)  
Unpublished (2001)  
Contact: Dr. Judith F. Margolin  
Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-824-4536  
Fax: 832-825-4038  
Email: clones@tccc.org  
Seq primer: M13 primer

## FEATURES

source  
1..488  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TCAAP6371"  
/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="myeloid cell"  
/dev\_stage="pediatric 6 years"  
/lab\_host="DH10B"  
/note="Vector: lambda pSB; Site\_1: BamHI; Site\_2: EcoRI;  
First strand cDNA was primed with an anchored  
XhoI-oligo(dT) primer [5'GGAGGATCGAGCGCCGAGGAGGAG(T)VN  
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand  
was primed with a BamHI-dC primer

15'AGAGAGCTCGGATCCGGCGCCCAATAATAAT(C) 3']  
Double-stranded cDNA was then digested with BamHI and XhoI  
and directionally cloned into the BamHI and SalI sites of  
lambda pSB vector. Library went through one round of  
normalization. Library was constructed by Wei Yu at RIKEN  
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,  
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,  
Schneider C, Hayashizaki Y, High efficiency selection of  
full-length cDNA by improved biotinylated cap trapper..  
DNA Res 4: 1, 61-6, Feb 28, 1997).

BASE COUNT 112 a 134 c 145 g 96 t  
ORIGIN

## Alignment Scores:

Pred. No.: 5,88e-89 Length: 488  
Score: 763.00 Matches: 152  
Percent Similarity: 97.48% Conservativeness: 3  
Best Local Similarity: 95.60% Mismatches: 4  
Query Match: 66.35% Indels: 1  
DB: 13 Gaps: 0

US-09-817-199B-2 (1-223) x BM149118 (1-488)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
|||||  
Db 12 ATGACGGGCACGCCAGCGCCGTTGCCACCGGGATGGCGAGGCCCGAGCGCTCCCGG 71

QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
|||||  
Db 72 CCCGTCAGTCCGAGCTTTT-CTCAGGGCAAGGTGCTCTCTGGGAGACACAGCGGTC 130

QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
|||||  
Db 131 GCAAAACATGTTCTCTGATCCAAATCAAGACGCGGCTTCTCTCGGAACTTCATA 190

QY 61 AlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
|||||  
Db 191 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGTGATGGTGGCGTGTGAGAGTGAG 250

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTrpTyr 100  
|||||  
Db 251 CTGCAGATCTGGACACGCTGNGCAGACAGCGTTCGAGCGGTACCCATGCTTATATAC 310

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsn 120  
|||||  
Db 311 AGAGATGTCAGGGCTTCTCTGTGTATGACATCAGCAACAAATCTTCTTCGACAAC 370

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140  
|||||  
Db 371 ATCAGGGCTGCTCCTCAGTATTCAGGATGATGCGGAGGAGCGTGTGATCATGCTGTG 430

QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159  
|||||  
Db 431 CTAGGCAACAGCGCGGATATGAGCAGCGCAAGAGTGTCCCTTCGCAAGACGAGAG 487

## RESULT 15

BQ687186  
LOCUS  
DEFINITION  
AGENCOURT\_8034867 NIH\_MGC\_110.Homo sapiens cDNA clone IMAGE:6206869  
5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
BQ687186  
BQ687186.1 GI:21812502  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 888)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabps@mail.nih.gov

## Tissue Procurement: ATCC

CNA Library Preparation: Rubin Laboratory  
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2362 row: a column: 14  
 High quality sequence stop: 662.

FEATURES  
source

Location/Qualifiers  
 1. .888  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:620689"  
 /clone\_lib="NIH\_MGC\_110"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

BASE COUNT 172 a 279 c 289 g 148 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 3.77e-88 Length: 888  
 Score: 760.00 Matches: 149  
 Percent Similarity: 76.62% Conservative: 28  
 Best Local Similarity: 64.50% Mismatches: 30  
 Query Match: 66.09% Indels: 24  
 DB: 14 Gaps: 3

US-09-817-199B-2 (1-223) x BQ687186 (1-888)

Qy 4 ThrProGlyAla-----ValAlaThrArgAspGlyGluAlaProGluArgSer----- 19  
 Db 81 ACCCCCGCTGCCTCCACGCTCCGCCACCGCCCAACGGCGCGCTCGGGACT 140  
 Qy 20 -----ProProCysSerPro----- 24  
 Db 141 GCGCTTTCGGCCCGCGCGCGCCCAACGGCGCGCTTGCAGCCGCGCGCCCTCGCTT 200  
 Qy 25 -----SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39  
 Db 201 GCGCGCGGTGCGACTTCTACGACGTCGCCCTTCAAGGTCTGCTGGTGGGGACTCGGGT 260  
 Qy 40 ValGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59  
 Db 261 GTGGGAAGACCTCTGCTGCTGCGATTCAGGATGCTGCTTCTCGCGGGGACCTTC 320  
 Qy 60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79  
 Db 321 ATCTCCACCGTAGGATGACTTCCGGAACAAGTTCTGGACGTGGATGGTGTGAAGGTG 380  
 Qy 80 LysLeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99  
 Db 381 AAGCTGCAGATGTGGACACAGTGTTCAGGCGGTTCCGACGTGTACCCATGCCTAC 440  
 Qy 100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAsp 119  
 Db 441 TACCGGATGCTCATGTCTGCTGCTGTCTACGATGTTCACCAACAGGCGCTCCTTTGAC 500  
 Qy 120 AsnIleArgAlaThrLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMet 139  
 Db 501 AACATCCAGGCGCTGGTCCAGATCCAGCCACGACGACGACGCGCTCATG 560  
 Qy 140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159  
 Db 561 CTGCTGGGGACAAAGGTGGACTCTGCCCATAGGCGTGTGGTGAAGAGGAGGACGGGAG 620

Qy 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179  
 Db 621 AAGTGGCCAAAGGAGTATGGACTGCCCTTCATGGAGACCGCCACAGCGGCTCAAC 680  
 Qy 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla 199  
 Db 681 GTGACTGGCCTTCACAGCCCTAGCAAGAGGTTGAAGACGCGCTCCCTGAAGGCTCCC 740  
 Qy 200 AspGluProSerPheGlnIleArgAspTyrVal 210  
 Db 741 AGCGAGCGCGCTTCCGCGCTGCATGATTACGTT 773

Search completed: June 23, 2003, 20:56:39  
 Job time : 1388 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model.

Run on: June 23, 2003, 15:56:08 ; Search time 228 Seconds  
(without alignments)  
2202.612 Million cell updates/sec

Title: US-09-817-199B-2

Perfect score: 1150

Sequence: 1 MGTGCAVATRDGEAPERSP.....FQIRDYVESQKRRSCCSFM 223

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPO.spool/US09817199/runat\_18062003\_145101\_785/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_101002 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPWT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817199 -CGN\_1\_1\_263 -runat\_18062003\_145101\_785 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 1150  | 100.0       | 1106   | 22    | Human polynucleoti |
| 2          | 1150  | 100.0       | 2323   | 22    | Human polynucleoti |
| 3          | 1145  | 99.6        | 1109   | 22    | Nucleotide sequenc |
| 4          | 1144  | 99.5        | 2623   | 22    | cDNA encoding nove |
| 5          | 1140  | 99.1        | 875    | 24    | Human Ras protein  |
| 6          | 1133  | 98.5        | 1316   | 22    | cDNA encoding nove |
| 7          | 1034  | 89.9        | 843    | 20    | DNA sequence encod |
| 8          | 969   | 84.3        | 576    | 24    | Human degranulatio |
| 9          | 942   | 81.9        | 576    | 24    | Mouse degranulatio |
| 10         | 770   | -           | 475    | 23    | DNA encoding novel |
| 11         | 731   | 63.6        | 1340   | 19    | Human RAB protein, |
| 12         | 697   | 60.6        | 2001   | 23    | Drosophila melanog |
| 13         | 535   | 46.5        | 964    | 22    | cDNA encoding nove |
| 14         | 516   | 44.9        | 911    | 21    | Arabidopsis thalia |
| 15         | 515.5 | 44.8        | 1129   | 21    | Zea mays DNA fragm |
| 16         | 515   | 44.8        | 1203   | 21    | Human ovarian anti |
| 17         | 514.5 | 44.7        | 2411   | 24    | Human rab8 homolog |
| 18         | 513.5 | 44.7        | 1986   | 22    | Human cancer agent |
| 19         | 511.5 | 44.5        | 2497   | 22    | Human cancer agent |
| 20         | 511.5 | 44.5        | 2497   | 22    | Human cancer agent |
| 21         | 511.5 | 44.5        | 2497   | 22    | Human cancer agent |
| 22         | 511.5 | 44.5        | 2497   | 23    | Human prostate exp |
| 23         | 511.5 | 44.5        | 2497   | 23    | Human prostate exp |
| 24         | 511.5 | 44.5        | 2497   | 23    | Drosophila melanog |
| 25         | 508.5 | 44.2        | 1540   | 23    | Arabidopsis thalia |
| 26         | 507   | 44.1        | 777    | 21    | Arabidopsis thalia |
| 27         | 507   | 44.1        | 881    | 24    | Arabidopsis thalia |
| 28         | 507   | 44.1        | 1023   | 21    | Arabidopsis thalia |
| 29         | 507   | 44.1        | 1025   | 21    | Arabidopsis thalia |
| 30         | 507   | 44.1        | 3077   | 22    | Human cDNA sequenc |
| 31         | 506.5 | 44.0        | 956    | 21    | Nucleotide sequenc |
| 32         | 506   | 44.0        | 674    | 23    | DNA encoding novel |
| 33         | 506   | 44.0        | 1537   | 21    | Human prostate can |
| 34         | 506   | 44.0        | 3533   | 22    | Human cDNA sequenc |
| 35         | 505.5 | 44.0        | 716    | 21    | Human Rab10 cDNA.  |
| 36         | 505.5 | 44.0        | 861    | 21    | Human Rab10 cDNA # |
| 37         | 503   | 43.7        | 959    | 21    | Arabidopsis thalia |
| 38         | 502   | 43.7        | 1101   | 21    | Arabidopsis thalia |
| 39         | 501   | 43.6        | 866    | 22    | Human cDNA clone ( |
| 40         | 501   | 43.6        | 1161   | 22    | Nucleotide sequenc |
| 41         | 498   | 43.3        | 888    | 21    | Canine Rab10 cDNA. |
| 42         | 497.5 | 43.3        | 932    | 21    | Arabidopsis thalia |
| 43         | 490   | 42.6        | 2247   | 23    | Drosophila melanog |
| 44         | 488   | 42.4        | 705    | 21    | Arabidopsis thalia |
| 45         | 483   | 42.0        | 666    | 21    | Arabidopsis thalia |

ALIGNMENTS

RESULT 1

AAI59066

ID AAI59066 standard; cDNA; 1106 BP.

XX AC AAI59066;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 1269.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia; sq.

XX OS Homo sapiens.

XX WO200153312-A1.  
 XX PD 26-JUL-2001.  
 XX PF 26-DEC-2000; 2000WO-US34263.  
 XX 21-JAN-2000; 2000US-0488725.  
 XX PR 25-APR-2000; 2000US-0552317.  
 XX PR 09-JUL-2000; 2000US-0598042.  
 XX PR 19-JUL-2000; 2000US-0620312.  
 XX PR 03-AUG-2000; 2000US-0653450.  
 XX PR 14-SEP-2000; 2000US-0662191.  
 XX PR 19-OCT-2000; 2000US-0693036.  
 XX PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 XX P-PSDB; AAM39910.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -  
 XX Claim 1; SEQ ID NO 1269; 10078pp; English.  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 XX the encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
 XX in gene therapy. A composition containing a polypeptide or polynucleotide  
 XX system, such as peripheral nervous injuries, peripheral neuropathy and  
 XX localised neuropathies and central nervous system diseases, such as  
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 XX utilisation of the activities such as: Immune system suppression,  
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 XX assays for receptor activity, arthritis and inflammation, leukaemias and  
 XX C.N.S disorders.  
 XX Note: The sequence data for this patent did not form part of the printed  
 XX specification.  
 XX SQ Sequence 1106 BP; 260 A; 307 C; 321 G; 218 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,36e-128 Length: 1106  
 Score: 1150.00 Matches: 223  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-817-199B-2 (1-223) x AAI59066 (1-1106)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 DB 28 ATGACGGGACGGCAGGCGCGTTCCTCCACCGGGATGCGAGGCGCCGCGGCTCCCG 87  
 QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 DB 88 CCTCAGTCCGAGTACGACTCAGCGGCAAGGTGATGCTCTGGGAGACACAGGCGTC 147  
 QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheTle 60  
 DB 148 GCGAAACATGTTCTCTGATCCATTCACACGGGCGCTTCCTCGGAACTTCATA 207  
 QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80

DB 208 GCCACCGTCGGCATAGACTTTCAGGAACAAGGTGGTGTGACTGTGGATGGCGCTGAGAGTGAAG 267  
 QY 81 LeuGlnIleTrrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 DB 268 CTGCAGATCTGGGACACCGCTGGCAGGAACGGTTCGGAAGCGTCAACCCATGCTATTATAC 327  
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 DB 328 AGAGATGCTCAGGCGCTTCTCTGTATGATCATCACCACAAATCTCTTTCGACAAAC 387  
 QY 121 IleArgAlaTrrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 DB 388 ATCAGGCGCTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447  
 QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 DB 448 CTAGGCAACAGCGGATATGAGCAGCGAAGAGTATCGTTCCTCCGACGAGGAGACG 507  
 QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 DB 508 TTGGCCAGGAGTACGGTGTTCCTCTCTGGAGACCGCCAGGACCTGAATACCGGCGGCGAT 567  
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 DB 568 GAGTTAGCTTCTGCGCATCCCAAGGACCTGAATACCGGCGGCGGCGATCAGCGGAT 627  
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 DB 628 GAGCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAGACGCGCCAGACTGCCATGTGC 687  
 QY 221 SerPheMet 223  
 DB 688 TCCTTTCATG 696  
 RESULT 2  
 AAI60852  
 ID AAI60852 standard; cDNA; 2323 BP.  
 AC AAI60852;  
 XX 22-OCT-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 4841.  
 XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 XX peripheral nervous system; neuropathy; central nervous system; CNS;  
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 XX leukaemia; ss.  
 XX Homo sapiens.  
 OS WO200153312-A1.  
 PN 26-JUL-2001.  
 PD 26-DEC-2000; 2000WO-US34263.  
 PF 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.  
 DR P-PSDB; AAM41696.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX Claim 1; SEQ ID NO 4841; 10078pp; English.  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369), and  
 CC the encoded polypeptides (AAM38642-AAAM42213), with neurotropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Alignment Scores:  
 Pred. NO.: 6,74e-128 Length: 2323  
 Score: 1150.00 Matches: 223  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-817-199B-2 (1-223) x AAI60852 (1-2323)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 Db 28 ATGACGGGACCGCCAGCGCGCTTCCACCCCGGATGGCGGCGCCGCGCGCTCCCG 87  
 Qy 21 ProCysSerProSerTyrrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 Db 88 CCTGCGAGTCGAGCTACGACCTCAGCGGGAAGGTGCTCTGGGAGACACAGGCGTC 147  
 Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 Db 148 GGCAAAACATGTTTCTTCGTATCCCAATTCAAAGACGGGCGCTTCTGTCCGGAACCTTCATA 207  
 Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
 Db 208 GCCACCTCGGCATAGACTTCAGAACCAAGGTGTGATGCGTGGATGCGTGAAG 267  
 Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrr 100  
 Db 268 CTGCAGATCTGGACACCGCTGGCAGAACGCTTCCGAGCGTCACCATGCTTATATAC 327  
 Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuTyrrAspIleThrAsnLysSerSerPheAspAsn 120  
 Db 328 AGAGATGCTCAGGCGCTTGTCTGCTATGACATCACCAACAAATCTCTCTTCGCAAC 387  
 Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrrAlaGlnArgAspValValIleMetLeu 140  
 Db 388 ATCAGGCGCTGGCTCACTAGATTCATGATGATGCCAGAGGGCGTGGTATCATCTG 447  
 Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 Db 448 CTAGGCACACAGCGGATATGAGCAGCAAGAGTGTATCGTTCGAGACGAGAGACC 507  
 Qy 161 LeuAlaArgGluTyrrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 Db 508 TTGGCCAGGGAGTACGCTGTCTCCCTCTGGAGACCAGCCCAAGAGTGGCATCAATGTG 567

Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrrArgAlaGlyHisGlnAlaAsp 200  
 Db 568 GAGTTAGCTTCTTGGCCATCGCCAAAGAACTGAAATACCGCGCGGCATCAGCGGAT 627  
 Qy 201 GluProSerPheGlnIleArgAspTyrrValGluSerGlnLysLysArgSerSerCysCys 220  
 Db 628 GAGCCAGCTTCCAGATCCGAGACTATGTAGATCCCAAGAAAGCGCTCCAGCTGCTGC 687  
 Qy 221 SerPheMet 223  
 Db 688 TCCTTCATG 696

RESULT 3  
 AAH75184  
 ID AAH75184 standard; DNA; 1109 BP.  
 XX AAH75184;  
 AC AAH75184;  
 XX 13-NOV-2001 (first entry)  
 DT Nucleotide sequence of human 32712 G-protein.  
 DE Human: G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;  
 XX congenital anomaly; pulmonary congestion; oedema; haemorrhage;  
 KW adult respiratory distress syndrome; Goodpasture's syndrome;  
 KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;  
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;  
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;  
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;  
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;  
 KW chronic bacterial meningoencephalitis; multiple sclerosis;  
 XX amytropic lateral sclerosis; stroke; Huntington's disease; ss.  
 OS Homo sapiens.

Key Location/Qualifiers  
 CDS 124..699  
 FT /\*tag= a  
 FT /product= "G-protein"  
 PN WO200164887-A2.  
 PD 07-SEP-2001.  
 XX 27-FEB-2001; 2001WO-US06292.  
 PF 29-FEB-2000; 2000US-0185606.  
 PR (MILL-) MILLENNIUM PHARM INC.  
 PA Meyers RA;  
 PI WPI: 2001-550182/61.  
 DR P-PSDB; AAG67156.  
 DR Novel human small G-protein polypeptides and polynucleotides for  
 XX treating lung disorders, liver disorders and brain disorders -  
 PS Claim 2; Fig 26; 151pp; English.

CC The present sequence encodes a human G-protein. The specification  
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The  
 CC G-protein polypeptides and polynucleotides are useful as a target for  
 CC diagnosis and treatment of G-protein mediated or related disorders,  
 CC and for identifying agonists and antagonists for diagnosis and  
 CC treatment. They are useful for treating disorders of lung (e.g.  
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory  
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,  
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver  
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,  
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and  
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute

CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic  
 CC bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral  
 CC sclerosis, stroke and Huntington's disease).  
 SQ Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 9,45e-128 | Length:       | 1109 |
| Score:                 | 1145.00   | Matches:      | 222  |
| Percent Similarity:    | 100.00%   | Conservative: | 0    |
| Best Local Similarity: | 100.00%   | Mismatches:   | 0    |
| Query Match:           | 99.57%    | Indels:       | 0    |
| DB:                    | 22        | Gaps:         | 0    |

US-09-817-199b-2 (1-223) x AAH75184 (1-1109)

|    |     |  |     |
|----|-----|--|-----|
| QY | 2   | ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro | 21  |
| DB | 31  | ACGGGCACGCCAGCGCCGTTGCCACCGGGATGGCGAGCCCGCGGCGTCCCGGCC       | 90  |
| QY | 22  | CysSerProSerThrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly | 41  |
| DB | 91  | TCGAGTCCGAGTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGC  | 150 |
| QY | 42  | LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla | 61  |
| DB | 151 | AAACATGTTCTCTGATCCAAATCAAGACGGGCTTCTCTCGGAACCTTCATAGCC       | 210 |
| QY | 62  | ThrValGlyLeuAspPheArgAsnLysValThrValAspGlyValArgValLysLeu    | 81  |
| DB | 211 | ACCGTCGGCATAGACTTCAAGAACAGGTGGTGAAGTGGTGGTGAAGTGAAGCTG       | 270 |
| QY | 82  | GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTrpArg    | 101 |
| DB | 271 | CAGATCTGGACACCGTGGCGAGGACGGTTCGGAAGCGTCACCCATGCTTATTACAGA    | 330 |
| QY | 102 | AspAlaGlnAlaLeuLeuLeuLeuTrpAspIleThrAsnLysSerSerPheAspAsnIle | 121 |
| DB | 331 | GATGCTCAGCGCTTCTCTGCTGCTGATGACATCACCAACAACTCTTTCGACACATC     | 390 |
| QY | 122 | ArgAlaTrpLeuThrGluIleHisGluTrpAlaGlnArgAspValValIleMetLeuLeu | 141 |
| DB | 391 | AGGGCTGGCTCAGTACGATTCATGATTCATGATGCCAGAGGGAGCTGATGCTGCTA     | 450 |
| QY | 142 | GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu | 161 |
| DB | 451 | GCACACAGCGGATATGAGCAGCAGGAGTATCCGTTCCGAAGACGGAGACCTTG        | 510 |
| QY | 162 | AlaArgGluTrpGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu | 181 |
| DB | 511 | GCAGGGAGTACGGTGTTCCTTCTGGAGACCGCCAGACTGGCATGAATGTGGAG        | 570 |
| QY | 182 | LeuAlaPheLeuAlaIleAlaLysGluLeuLysTrpArgAlaGlyHisGlnAlaAspGlu | 201 |
| DB | 571 | TTAGCTTTCGGCCATCCCAAGGAATGAATACCGGGCGGCGCATCAGCGCGATAG       | 630 |
| QY | 202 | ProSerPheGlnIleArgAspTrpValGluSerGlnLysLysArgSerSerCysCysSer | 221 |
| DB | 631 | CCCAGTCCAGATCCGAGACTATGTAGTCTCCAGAGAGAGCGCTCCAGTGTGCTTCC     | 690 |
| QY | 222 | PheMet   | 223 |
| DB | 691 | TTCATG   | 696 |

# RESULT 4

AAAS27036  
 ID AAAS27036 standard; cDNA; 2623 BP.

XX AC AAAS27036;

XX DT 07-NOV-2001 (first entry)

XX DE cDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
 KW organ transplant rejection; infection; hepatitis C; blood disorder;  
 KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
 KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
 KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;  
 XX acquired immune deficiency syndrome.  
 OS Homo sapiens.  
 XX  
 PN WO200154733-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01312.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 09-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.



PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 14-SEP-2000; 2000US-0233066.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 21-SEP-2000; 2000US-0234225.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236328.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX P-PSDB; AAU17119.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders  
XX  
XX Claim 1; SEQ ID No 71; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative  
XX disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
XX abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
XX disorders (e.g. glomerulonephritis), cardiovascular disorders  
XX (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
XX wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
XX Addison's disease), reproductive system disorders, gastrointestinal  
XX disorder (inflammatory disorders), liver disorders (cirrhosis),  
XX as stimulators of B-cell responsiveness to pathogens, activators of  
XX T-cells, to induce higher affinity antibodies, and as a means to induce  
XX tumour proliferation in pathologies e.g. acquired immune deficiency  
XX syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
XX pathway protein coding sequences and PCR primers of the invention.

## Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 4,22e-127 | Length:       | 2623 |
| Score:                 | 1144.00   | Matches:      | 222  |
| Percent Similarity:    | 99.55%    | Conservative: | 0    |
| Best Local Similarity: | 99.55%    | Mismatches:   | 1    |
| Query Match:           | 99.48%    | Indels:       | 0    |
| DB:                    | 22        | Gaps:         | 0    |

US-09-817-199b-2 (1-223) x AAS27036 (1-2623)

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro | 20  |
| DB | 12  | ATGACGGGCACGGCAGCGCCGCTTGCACCCGGGATGGCGAGGCCGCCGAGCTCCCG     | 71  |
| QY | 21  | ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal | 40  |
| DB | 72  | CCCTGCAGTCCGAGCTAGGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC | 131 |
| QY | 41  | GlyIysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle | 60  |
| DB | 132 | GGCAAAACATGTTTCTGTGATCCAAATCAAGACGGGGCGCTTCTGTCGGAAGCTTCATA  | 191 |
| QY | 61  | AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys | 80  |
| DB | 192 | GCCACCGTCGGCATAGACTTTCAGCAACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  | 251 |

QY 81 LeuGlnIleTTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 Db |||||||  
 252 CTGCAGATCTGGACACCGCTGGGCAGGAACGGTTCGAAGCGTCACCCATGCTTATTAC 311  
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
 Db |||||||  
 312 AGAGATGCTCAGGCGCTTCTGCTGTATGACATCACCACAACTCTCTTCGCAAC 371  
 QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 Db |||||||  
 372 ATCAGGGCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 431  
 QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 Db |||||||  
 432 CTAGCCACACAGCGGATATGAGCAGCAAGAGTGTCCGTCGAAGACGGAGAGACC 491  
 QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 Db |||||||  
 492 TTGGCCAGGGAGTACGGTGTCTCCCTTCTGGAGACCAGCCCAAGACTGGCATGAATGTG 551  
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 Db |||||||  
 552 GAGTTAGCCTTCTGGCCATCGCCAAAGAACTGAAATACCGGGCGGCATCAGCGCAT 611  
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 Db |||||||  
 612 GAGCCAGCTTCCAGATCCGAGACTATGATGATGATGATGATGATGATGATGATGATG 671  
 QY 221 SerPheMet 223  
 Db |||||||  
 672 TCCTTCATG 680

## RESULT 5

ID ABN83696  
 XX ABN83696 standard; cDNA; 875 BP.  
 AC ABN83696;  
 XX  
 DT 27-AUG-2002 (first entry)  
 DE Human Ras protein 3 (RASP-3) cDNA.  
 XX  
 KW Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;  
 KW immunosuppressive; antiinflammatory; signal transduction; gene;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 93..668  
 FT /\*tag= a  
 FT /product= "RASP-3"  
 XX  
 PN US6391580-B1.  
 XX  
 PD 21-MAY-2002.  
 XX  
 PF 08-MAY-1998; 98US-0075454.  
 XX  
 PR 12-DEC-1996; 96US-0766551.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Hillman JL, Tang YT, Lal P, Guegler KJ, Corley NC, Patterson C;  
 PI Batra S, Baughn MR;  
 XX  
 DR WPI; 2002-498774/53.  
 DR P-PSDB; ABB76426.  
 XX  
 PT New isolated polynucleotides encoding Ras proteins designated RASP-1  
 PT and RASP-4, for diagnosing, preventing and treating disorders  
 PT associated with cell proliferation, particularly cancer and immune

## disorders

PT  
 PS Example; Column 57-60; 34pp; English.  
 XX

CC The present sequence is that of a cDNA clone encoding novel human  
 CC Ras protein 3 (RASP-3, see ABB76426). Nucleic acids encoding  
 CC RASP-3 were initially identified in Incyte Clone 1528559 from a  
 CC mononuclear cell cDNA library (UMCL5701) using a computer search  
 CC for amino acid sequence alignments. The present consensus  
 CC sequence was derived from overlapping and/or extended nucleic  
 CC acids in Incyte clone 1528559 and shot-gun sequence SAEA03135,  
 CC SAE10396 and SABC10855. RASP-3 shows homology to rat Rab26,  
 CC Northern analysis showed expression of RASP-3 in haematopoietic  
 CC and immunological cDNA libraries, all associated with inflammation  
 CC and the immune response. The invention provides 7 novel human Ras  
 CC proteins (RASP-1 to -7) and polynucleotides, expression vectors,  
 CC host cells, antibodies, agonists and antagonists. It also provides  
 CC methods for diagnosing, treating or preventing disorders associated  
 CC with RASP expression, especially cancer and immune disorders. A  
 CC fragment of the present sequence, from about nucleotide 92 to about  
 CC nucleotide 153, is useful as a hybridisation probe.

XX  
 SQ Sequence 875 BP; 203 A; 247 C; 260 G; 165 T; 0 other;

## Alignment Scores:

Pred. No.: 2,7e-127 Length: 875  
 Score: 1140.00 Matches: 221  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.13% Indels: 0  
 DB: 24 Gaps: 0

US-09-817-199B-2 (1-223) x ABN83696 (1-875)

QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProCys 22  
 Db |||||||  
 3 GGCAGCCAGCGCGGTGTCACCCGGATGGCGAGGCCCGGAGCGTCCCGCCCTGC 62  
 QY 23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42  
 Db |||||||  
 63 AGTCGAGCTACGACCTCACGGCAAGGTGTGTTCTGGAGACACACAGCGCTCGCAAA 122  
 QY 43 ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleLeuThr 62  
 Db |||||||  
 123 ACATGTTCTGTATCCAAATTCAGAGACGGGGCTTCCTGTCGGAACCTTCATAGCCACC 182  
 QY 63 ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82  
 Db |||||||  
 183 GTCGCATAGACTTCAGGAACAAGTGTGCTGACTGTGGATGGGTGAGAGTGAAGCTGCAG 242  
 QY 83 IleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102  
 Db |||||||  
 243 ATCTGGACACCGCTGGCAGGAACGGTTCGGAAGCGTCACCCATGCTTATTACAGAGAT 302  
 QY 103 AlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122  
 Db |||||||  
 303 GCTCAGGCTTGTCTGCTGTATGACATCACCAACAAATCTCTTCGCAACATCAGG 362  
 QY 123 AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142  
 Db |||||||  
 363 GCCTGGCTCACTGAGATTCATGATATGCCAGAGGGACGGTGTGATCATCTGCTGATGCC 422  
 QY 143 AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrIleAla 162  
 Db |||||||  
 423 AACAGGCGGATATGAGCAGCGAAAGTGTATCGTTCGGAAGACGGAGAGACCTTGGCC 482  
 QY 163 ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182  
 Db |||||||  
 483 AGGGAGTACGGTGTCTCCCTTCTGGAGACCAAGACTGGCATGAATGTGGAGTTA 542  
 QY 183 AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro 202  
 Db |||||||  
 543 GCCTTCTGCGCCATCGCAAGAACTGAATACCGGGCGGCATCAGCGCGATGAGGCC 602

|          |                              |   |     |
|----------|------------------------------|---|-----|
| Qy       | 203                          | SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSerPhe              | 222 |
| Db       | 603                          | AGCTTCCAGATCCGAGACTATGTAGAGTCCGAGAAGAAGGCTCCAGCTGCTGCTCTTC                | 662 |
| Qy       | 223                          | Met 223   |     |
| Db       | 563                          | ATG 665   |     |
| RESULT 6 |                              |   |     |
| AAS27458 |                              |   |     |
| ID       | AAS27458                     | standard; cDNA; 1316 BP.  |     |
| AC       | AAS27458;                    |   |     |
| XX       | 07-NOV-2001                  | (first entry)   |     |
| DT       |                              |   |     |
| XX       |                              |   |     |
| DE       |                              | cDNA encoding novel signal transduction pathway protein, Seq ID 493.      |     |
| XX       |                              |   |     |
| KW       |                              | Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;   |     |
| KW       |                              | antiinflammatory; anti-Hiv; antibacterial; antiinflammatory; cancer;      |     |
| KW       |                              | immune system disorder; rheumatoid arthritis; inflammatory condition;     |     |
| KW       |                              | organ transplant rejection; infection; hepatitis C; blood disorder;       |     |
| KW       |                              | sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;      |     |
| KW       |                              | neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;     |     |
| KW       |                              | chromosomal abnormality; Down syndrome; ischaemia; renal disorder;        |     |
| KW       |                              | cardiovascular; respiratory; wound healing; endocrine; Addison's disease; |     |
| KW       |                              | reproductive system; gastrointestinal; liver disorder; AIDS; ss;          |     |
| XX       |                              | acquired immune deficiency syndrome.                                      |     |
| XX       |                              |   |     |
| OS       |                              | homo sapiens.   |     |
| XX       |                              |   |     |
| PN       | WO200154733-A1.              |   |     |
| XX       |                              |   |     |
| PD       | 02-AUG-2001.                 |   |     |
| XX       |                              |   |     |
| PF       | 17-JAN-2001; 2001WO-US01312. |   |     |
| XX       |                              |   |     |
| PR       | 31-JAN-2000; 2000US-0179065. |   |     |
| PR       | 04-FEB-2000; 2000US-0180628. |   |     |
| PR       | 24-FEB-2000; 2000US-0184664. |   |     |
| PR       | 02-MAR-2000; 2000US-0186350. |   |     |
| PR       | 16-MAR-2000; 2000US-0189874. |   |     |
| PR       | 17-MAR-2000; 2000US-0190076. |   |     |
| PR       | 18-APR-2000; 2000US-0198123. |   |     |
| PR       | 19-MAY-2000; 2000US-0205515. |   |     |
| PR       | 07-JUN-2000; 2000US-0209467. |   |     |
| PR       | 28-JUN-2000; 2000US-0214886. |   |     |
| PR       | 30-JUN-2000; 2000US-0215135. |   |     |
| PR       | 07-JUL-2000; 2000US-0216647. |   |     |
| PR       | 07-JUL-2000; 2000US-0216880. |   |     |
| PR       | 11-JUL-2000; 2000US-0217487. |   |     |
| PR       | 11-JUL-2000; 2000US-0217496. |   |     |
| PR       | 14-JUL-2000; 2000US-0218290. |   |     |
| PR       | 26-JUL-2000; 2000US-0220963. |   |     |
| PR       | 26-JUL-2000; 2000US-0220964. |   |     |
| PR       | 14-AUG-2000; 2000US-0224518. |   |     |
| PR       | 14-AUG-2000; 2000US-0224519. |   |     |
| PR       | 14-AUG-2000; 2000US-0225213. |   |     |
| PR       | 14-AUG-2000; 2000US-0225214. |   |     |
| PR       | 14-AUG-2000; 2000US-0225266. |   |     |
| PR       | 14-AUG-2000; 2000US-0225267. |   |     |
| PR       | 14-AUG-2000; 2000US-0225268. |   |     |
| PR       | 14-AUG-2000; 2000US-0225270. |   |     |
| PR       | 14-AUG-2000; 2000US-0225447. |   |     |
| PR       | 14-AUG-2000; 2000US-0225757. |   |     |
| PR       | 14-AUG-2000; 2000US-0225758. |   |     |
| PR       | 14-AUG-2000; 2000US-0225759. |   |     |
| PR       | 18-AUG-2000; 2000US-0226279. |   |     |
| PR       | 22-AUG-2000; 2000US-0226681. |   |     |
| PR       | 22-AUG-2000; 2000US-0226686. |   |     |
| PR       | 22-AUG-2000; 2000US-0227182. |   |     |
| PR       | 23-AUG-2000; 2000US-0227009. |   |     |

PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-465460/50.  
 DR P-PSDB; AAU17541.  
 XX  
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders -  
 PS Claim 1; SEQ ID No 493; 880pp; English.  
 XX  
 XX The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
 CC pathway protein coding sequences and PCR primers of the invention.  
 XX

Alignment Scores:  
 Pred. No.: 3,34e-126 Length: 1316  
 Score: 1133.00 Matches: 220  
 Percent Similarity: 99.10% Conservative: 0  
 Best Local Similarity: 99.10% Mismatches: 2  
 Query Match: 98.52% Indels: 0  
 DB: 22 Gaps: 0

US-09-817-199b-2 (1-223) x AAS27458 (1-1316)  
 QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21  
 |||||  
 Db 1 ACGGCGACGCCAGCGCGGTTGCCACCGGGATGGGAGGCCCGCCGAGCGCTCCCGGCC 60  
 |||||  
 QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41  
 |||||  
 Db 61 TGCAGTCGGAGCTACGACCTCACGGCAAGGTGATGCTCTCTGGGAGACACAGGCGTCCGC 120  
 |||||  
 QY 42 LysThrCysPheLeuIleGlnPheIysAspGlyAlaPheLeuSerGlyThrPheIleIle 61  
 |||||  
 Db 121 AAAACATGTTTCCTGATCCCAATCAAGACACGGGCTTCCTGTCGGGAACCTTCATAGCC 180  
 |||||  
 QY 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValValVal 81  
 |||||  
 Db 181 ACCGTGGCATAGACTTCAGSACAAAGGTGGTGTACTGTGGATGGCGTGGAGTAGTGAGCTG 240  
 |||||  
 QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101  
 |||||  
 Db 241 CAGATCTGGGACACCGCTGGGCGAGCAAGGTTCGGAACGCTCACCCATGCTTATTACAGA 300  
 |||||  
 QY 102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121  
 |||||  
 Db 301 GATGCTCAGGCTTGTCTGTGTATGATCATCACCAACAAATCTCTTCGACAACATC 360  
 |||||  
 QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141  
 |||||  
 Db 361 AGGCGCTGGCTCACTGAGATTGATGATGCCAGAGGGAGCTGGTGATCATGCTGCTA 420  
 |||||  
 QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161  
 |||||  
 Db 421 GGCAACAAGCGGATATGAGCAGCAAGAGTGTATCCGTCGGAAGACGAGAGACCTTG 480  
 |||||  
 QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181  
 |||||  
 Db 481 GCCAGGAGTACGGKGTTCCTTCTGGAGACCGCCAGACCTGGCATGAATGTGGAG 540  
 |||||  
 QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201  
 |||||  
 Db 541 TTAGCCTTCTGGCCATCGCCCAAGGAAGTGAATACCGCGCGGCATCAGCGCGATGAG 600  
 |||||  
 QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221  
 |||||  
 Db 601 CCCAGTTCAGATCCGAGACTATGATAGTCCCAAGAGAGCGCTCCAGCTCCAGCTGCTCC 660  
 |||||  
 QY 222 PheMet 223  
 |||||  
 Db 661 TTCATG 666  
 |||||  
 RESULT 7  
 AAX86720  
 ID AAX86720 standard; DNA; 843 BP.  
 XX  
 AC AAX86720;  
 XX  
 DT 27-OCT-1999 (first entry)  
 XX  
 DE DNA sequence encoding an exocytotic protein designated Exo2.  
 XX  
 KW Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;  
 KW inflammation; ss.  
 XX  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FH 124..699  
 FT CDS /tag= a  
 FT /transl\_except= (pos: 136..138, aa: xaa)  
 FT /transl\_except= (pos: 202..204, aa: xaa)  
 FT /note= "Xaa is an unspecified amino acid"  
 XX



Score: 969.00 Matches: 189  
 Percent Similarity: 100.00% Conservatives: 2  
 Best Local Similarity: 98.95% Mismatches: 0  
 Query Match: 84.26% Indels: 0  
 DB: 24 Gaps: 0

US-09-817-199B-2 (1-223) x ABA02774 (1-576)

QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52  
 DB 1 ATGCTTCTTGGAGACTCGGGCTCGGCAAAACACTGTTCTCTGATCAATTCAGAGCGG 60  
 QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72  
 DB 61 GCCTTCCTTCGGAACCTTATAGCCACCGCTGAGTATTCAGGAAAGGTGTG 120  
 QY 73 ThrValAspGlyValArgValLysLeuGlnIleTyrPheAlaGlyGlnGluArgPhe 92  
 DB 121 ACTGTGGATGGCTGAGAGTGAAGCTGAGATCTGGACACCGCTGGCAGGAACGGTTC 180  
 QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112  
 DB 181 CGAAGCGTCACCATGCTTATTACAGAGATGCTCAGGCTTCTCTGCTGTATGACATC 240  
 QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTyrLeuThrGluIleHisGluTyrAla 132  
 DB 241 ACCAACAAATCTCTTCGACAAACATCAGGGCTGGCTCAGTATTCATGATGATGATGCC 300  
 QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152  
 DB 301 CAGAGGGAGCTGGTATCATGCTGTAGCAACAGCGGATATGAGCAGCAAGAGATG 360  
 QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValIlePheLeuGluThr 172  
 DB 361 ATCCGTTCGGAAGCGGAGAGACCTTGGCCAGGAGTACGGTGTCCCTTCTCGAGACC 420  
 QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192  
 DB 421 AGCGCCAGACCTGGCATGATGAGTGTAGCTTCTGGCCATCGCCAAAGAACTGAAA 480  
 QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212  
 DB 481 TACCGGGCGGCGCATCAGCGCATGAGCCACCTCCAGATCCGAGACTATGTAGATGCC 540

# RESULT 9

ABA02773  
 ID ABA02773 standard; DNA; 576 BP.  
 AC ABA02773;  
 DT 07-FEB-2002 (first entry)  
 DE Mouse degranulation regulator encoding DNA SEQ ID NO 2.  
 KW Degranulation; mast cell; human; mouse; anti-allergic; ds.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 CDS 1..576  
 FT /tag= a  
 FT /transl\_except= (pos:247..249,aa:Lys)  
 FT /product= "degranulation regulator"

WO200179478-A1.

25-OCT-2001.

16-APR-2001; 2001WO-JP03268.

XX 19-APR-2000; 2000JP-0118408.  
 XX (DAIN ) DAINIPPON PHARM CO LTD.  
 XX Yamada T, Ido M;  
 PI WPI; 2002-041335/05.  
 DR P-PSDB; AAM52189.  
 XX Mast cell degranulation controller for treatment of allergies  
 PT Claim 10; Page 63-64; 85pp; Japanese.  
 PS The invention relates to a protein for regulating degranulation of mast  
 CC cells (degranulation regulators) and the encoding polynucleotides, with  
 CC anti-allergic activity, used in the treatment of allergies associated  
 CC with degranulation of mast cells.  
 XX Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

## Alignment Scores:

Pred. No.: 9.69e-104 Length: 576  
 Score: 942.00 Matches: 183  
 Percent Similarity: 98.43% Conservatives: 5  
 Best Local Similarity: 95.81% Mismatches: 3  
 Query Match: 81.91% Indels: 0  
 DB: 24 Gaps: 0

US-09-817-199B-2 (1-223) x ABA02773 (1-576)

QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52  
 DB 1 ATGCTTCTTGGAGACTCGGGCTCGGCAAAACACTGTTCTCTGATCAATTCAGAGCGG 60  
 QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72  
 DB 61 GCCTTCCTTCGGAACCTTATAGCCACCGCTGAGTATTCAGGAAAGGTGTG 120  
 QY 73 ThrValAspGlyValArgValLysLeuGlnIleTyrPheAlaGlyGlnGluArgPhe 92  
 DB 121 ACAGTGGATGGTGGCCAGGGTGAAGCTTCAGATCTGGACACTGAGGACAGAGCGCTTT 180  
 QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112  
 DB 181 CGCAGTGTGACCATGCTTATTACGAGATGCTCAGGCTTCTCTGTTGTATGACATC 240  
 QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTyrLeuThrGluIleHisGluTyrAla 132  
 DB 241 ACCAACCAATCTCTTCGACAAACATCAGGGCTGGCTCAGAGATTCATGATGATGCC 300  
 QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152  
 DB 301 CAGAGAGAGCTGGTATGATGCTTCTAGCAACAGCGGATGTAAGCAGCAAGAGGTG 360  
 QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValIlePheLeuGluThr 172  
 DB 361 ATCCGTTCGGAAGCTGGAGAGACACTGGCCAGGGAATATGTTCTCTTCATGAGACC 420  
 QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192  
 DB 421 AGTGCACAGACTGGCATGATGAGTGTGGCTTCTGGCAATTCAGGAACTGAGAA 480  
 QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212  
 DB 481 TACCGTTCGAGGAGGAGCGCTGATGAGCCACCTCCAGATCCGAGACTATGTGAGTCC 540  
 QY 213 GlnLysLysArgSerSerCysCysSerPheMet 223  
 DB 541 CAGAAGAGCGCTCCAGCTGCTGCTCTTTGTS 573

RESULT 10  
 AAS81431







```
Db 1477 TATGTCCGAGCATTTCTCCTCACACAGTTGGCATTTAGTATTTAGGAACAAGTGGTGTC 1536
Qy 74 ValAspGlyValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArg 93
Db 1537 GTCTGAGGACGGCGCTCAAGCTGCAATCTGGGACACAGCTGTCAGGAGCGATTCCGG 1596
Qy 94 SerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThr 113
Db 1597 AGCGTTACCCACGCCCTATTATCGGAGCGCACGCTCTACTGCTGTGTACGACGTGACC 1656
Qy 114 AsnLysSerPheAspAsnIleArgAlaTyrPleuThrGluIleHisGluTyrAlaGln 133
Db 1657 AACAGACACCATATGACACATTCGGCTGTGGCGGAGATCCGGAGTACGGCCAG 1716
Qy 134 ArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSer---SerGluArgVal 152
Db 1717 GAGGAGCTGTCTATCGCTTTTATAGCAACAGCGGACATGCAGCGGCGAGCGGCGAG 1776
Qy 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 1777 GTGAAGCGGAGGATGGGAGCGTTTGGGCGGGAGCACAACTGCTCCCTTCATGGAGACC 1836
Qy 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 1837 TCGGCCAAGACGGAGCTCATCTGGAGCTGCTTCACAGCGGTGCCAGGCAACTAAG 1896
Qy 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 1897 AGTCGGGGCTACGAGCAGCGGATGATGGAAGTTCAATGTGCATGATTTGTGCGGTGAC 1956
Qy 213 GlnLysLysArgSerSerCysCysSer 221
Db 1957 AATACAAGCGCGCTCGTTTGGCCC 1983

RESULT 13
AAS27040
ID AAS27040 standard; cdna; 964 BP.
XX
AC AAS27040;
XX
DT 07-NOV-2001 (first entry)
XX
DE cdna encoding novel signal transduction pathway protein, Seq ID 75.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
```



|    |                              |                 |               |
|----|------------------------------|-----------------|---------------|
| XX | 06-SEP-2000.                 |                 |               |
| PD |                              |                 |               |
| XX | 25-FEB-2000; 2000EP-0301439. | PR 02-JUL-1999; | 99US-0142055. |
| XX |                              | PR 06-JUL-1999; | 99US-0142390. |
| PF |                              | PR 08-JUL-1999; | 99US-0142803. |
| XX |                              | PR 09-JUL-1999; | 99US-0142920. |
| XX |                              | PR 12-JUL-1999; | 99US-0142977. |
| PR |                              | PR 13-JUL-1999; | 99US-0143542. |
| PR |                              | PR 14-JUL-1999; | 99US-0143624. |
| PR |                              | PR 15-JUL-1999; | 99US-0144005. |
| PR |                              | PR 16-JUL-1999; | 99US-0144085. |
| PR |                              | PR 16-JUL-1999; | 99US-0144086. |
| PR |                              | PR 19-JUL-1999; | 99US-0144325. |
| PR |                              | PR 19-JUL-1999; | 99US-0144331. |
| PR |                              | PR 19-JUL-1999; | 99US-0144332. |
| PR |                              | PR 19-JUL-1999; | 99US-0144333. |
| PR |                              | PR 19-JUL-1999; | 99US-0144334. |
| PR |                              | PR 20-JUL-1999; | 99US-0144335. |
| PR |                              | PR 20-JUL-1999; | 99US-0144352. |
| PR |                              | PR 20-JUL-1999; | 99US-0144632. |
| PR |                              | PR 20-JUL-1999; | 99US-0144884. |
| PR |                              | PR 21-JUL-1999; | 99US-0144814. |
| PR |                              | PR 21-JUL-1999; | 99US-0145086. |
| PR |                              | PR 21-JUL-1999; | 99US-0145088. |
| PR |                              | PR 22-JUL-1999; | 99US-0145085. |
| PR |                              | PR 22-JUL-1999; | 99US-0145087. |
| PR |                              | PR 22-JUL-1999; | 99US-0145089. |
| PR |                              | PR 22-JUL-1999; | 99US-0145192. |
| PR |                              | PR 23-JUL-1999; | 99US-0145145. |
| PR |                              | PR 23-JUL-1999; | 99US-0145218. |
| PR |                              | PR 23-JUL-1999; | 99US-0145224. |
| PR |                              | PR 26-JUL-1999; | 99US-0145276. |
| PR |                              | PR 27-JUL-1999; | 99US-0145913. |
| PR |                              | PR 27-JUL-1999; | 99US-0145918. |
| PR |                              | PR 27-JUL-1999; | 99US-0145919. |
| PR |                              | PR 28-JUL-1999; | 99US-0145951. |
| PR |                              | PR 02-AUG-1999; | 99US-0146386. |
| PR |                              | PR 02-AUG-1999; | 99US-0146388. |
| PR |                              | PR 02-AUG-1999; | 99US-0146389. |
| PR |                              | PR 03-AUG-1999; | 99US-0147038. |
| PR |                              | PR 04-AUG-1999; | 99US-0147204. |
| PR |                              | PR 05-AUG-1999; | 99US-0147302. |
| PR |                              | PR 05-AUG-1999; | 99US-0147192. |
| PR |                              | PR 06-AUG-1999; | 99US-0147260. |
| PR |                              | PR 06-AUG-1999; | 99US-0147303. |
| PR |                              | PR 06-AUG-1999; | 99US-0147416. |
| PR |                              | PR 09-AUG-1999; | 99US-0147493. |
| PR |                              | PR 09-AUG-1999; | 99US-0147935. |
| PR |                              | PR 10-AUG-1999; | 99US-0148171. |
| PR |                              | PR 11-AUG-1999; | 99US-0148319. |
| PR |                              | PR 12-AUG-1999; | 99US-0148341. |
| PR |                              | PR 13-AUG-1999; | 99US-0148585. |
| PR |                              | PR 13-AUG-1999; | 99US-0148684. |
| PR |                              | PR 16-AUG-1999; | 99US-0149368. |
| PR |                              | PR 17-AUG-1999; | 99US-0149175. |
| PR |                              | PR 18-AUG-1999; | 99US-0149426. |
| PR |                              | PR 20-AUG-1999; | 99US-0149722. |
| PR |                              | PR 20-AUG-1999; | 99US-0149723. |
| PR |                              | PR 20-AUG-1999; | 99US-0149929. |
| PR |                              | PR 23-AUG-1999; | 99US-0149902. |
| PR |                              | PR 23-AUG-1999; | 99US-0149930. |
| PR |                              | PR 25-AUG-1999; | 99US-0150566. |
| PR |                              | PR 26-AUG-1999; | 99US-0150884. |
| PR |                              | PR 27-AUG-1999; | 99US-0151085. |
| PR |                              | PR 27-AUG-1999; | 99US-0151086. |
| PR |                              | PR 27-AUG-1999; | 99US-0151080. |
| PR |                              | PR 30-AUG-1999; | 99US-0151303. |
| PR |                              | PR 31-AUG-1999; | 99US-0151438. |
| PR |                              | PR 01-SEP-1999; | 99US-0151930. |
| PR |                              | PR 07-SEP-1999; | 99US-0152363. |
| PR |                              | PR 10-SEP-1999; | 99US-0153070. |
| PR |                              | PR 13-SEP-1999; | 99US-0153758. |
| PR |                              | PR 15-SEP-1999; | 99US-0154018. |
| PR |                              | PR 16-SEP-1999; | 99US-0154039. |
| PR |                              | PR 20-SEP-1999; | 99US-0154779. |

```

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 3.18e-52
Score: 516.00
Percent Similarity: 66.52%
Best Local Similarity: 45.92%
Query Match: 44.87%
DB: 21
Length: 911
Matches: 107
Conservative: 48
Mismatch: 58
Indels: 20
Gaps: 6

US-09-817-199b-2 (1-223) x AAC38429 (1-911)
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
DB 28 CTAACATCTCTCCGATGCGGTT-----GCGCGGCAAGAGCT--- 66
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
DB 67 -----CGTTCAGACTATGATATCTCATCAGCTTCCTTCATCGGTGATAGCGGTG 120
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
DB 121 GGAAGAAAGTTGCTTCTACTGCGTTCGATGATGATCTTTCACCAAGC---AGTTTCATT 177
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
DB 178 ACTACATGGATGGATTCAGATCAAGATAGACAGTCGACTTGTGGGAAGCGTATCAAA 237
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
DB 238 TTGCAGATATGGACACTGCTGACAAAGACGTTTTCAGACTATAACCCACACATATAC 297
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
DB 298 AGAGGAGCGGATATATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 357

```

```

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
DB 358 ATTAGGAACATGGATGAAACAAACATTGACGAGCATGCGTCCGATAGTGTCAACAAGATATG 417
QY 141 LeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSerGluAspGlyGlu 159
DB 418 GTCGGTAAACAAAGCCGACATGGACGAAAGAAAGGGGTGCTCCNACATCAAAAGGACAA 477
QY 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
DB 478 GCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
QY 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGly 196
DB 538 CTCGAGCAGAGGTTTCTCTTCTATGCTAAAGACATAAAACAAGACTCACAGAAAGCGAT 597
QY 197 HisGlnAlaAspGluProSerPheGlnIle-----ArgAspTyrVal 210
DB 598 ACAAGCCGAGCCCAAGGATCAGATCACTAACAAAGATGCTAACAAAGCCTCATCG 657
QY 211 GluSerGlnLysLysArgSerSerCysCysSerPheMet 223
DB 658 TCTTCTACAAATGAGAAATCAGCTTGTGCTGAGTTATGTT 696

RESULT 15
AAC44482
ID AAC44482 standard; DNA; 1129 BP.
XX
AC AAC44482;
XX
DT 18-OCT-2000 (first entry)
DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

```

PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147433.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149388.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155119.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

|                        |          |               |      |
|------------------------|----------|---------------|------|
| Pred. No.:             | 4,95e-52 | Length:       | 1129 |
| Score:                 | 515.50   | Matches:      | 104  |
| Percent Similarity:    | 66.38%   | Conservative: | 50   |
| Best Local Similarity: | 44.83%   | Mismatches:   | 61   |
| Query Match:           | 44.83%   | Indels:       | 17   |
| DB:                    | 21       | Gaps:         | 6    |

US-09-817-199b-2 (1-223) x AAC44482 (1-1129)

```
QY      5  ProGlyAlaValAlaThrArgAspGlyGluAlaProGluArg----- 18
Db      127 CGGGACCGCGGGGTCGCGCGCGGTAGGAGGAGGAGGAGGCGGTATGGCG 186
QY      19  SerProPro-----CysSerProSerTyArgPheLeuThrGlyLysValMetLeuGly 36
Db      187 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY      37  AspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSer 56
Db      247 GACAGCGGTGTGGCAAGAGTTCCTGCTGTGGCGTCTCTGATGTTCTCTTCTACATA 306
QY      57  GlyThrPheLeuAlaThrValGlyLysPheArgAsnLysValValThrValAspGly 76
Db      307 ---AGCTTTATACCAATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 363
QY      77  ValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThr 96
Db      364 AAGCGTATAAAGCTACAGATTGGGTACGGCGGCGCGCGCGCGCGCGCGCGCGCGCG 423
QY      97  HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 116
Db      424 ACCGGGTACTACCGAGGAGCTATGGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
QY      117 SerPheAspAsnIleArgAlaThrPheLeuThrGluIleHisGluTyrAlaGlnArgVal 136
Db      484 TCTTTTACACATCCGAACTGGATCGGAACATGGAACATGGAACATGGAACATGGAACAT 543
QY      137 ValIleMetLeuLeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSer 155
Db      544 AACAAAGATTGGTTGGCAACAAAGCGCGATATGGAGCAAGCAAGAGGCGGTACTACT 603
QY      156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValPropheLeuGluThrSerAlaLys 175
Db      604 GCAAAAGGACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY      176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
Db      664 ACAACCTGATGTGGACCAAGTGTTCCTCCATCGCGCGCGCGCGCGCGCGCGCGCGCG 723
QY      196 GlyHisGlnAlaAspGluPro-----SerPheGlnIleArgAspTyrValGluSer 212
Db      724 GCGGAGACTGATTCAAGCCAGAGGACAAAGCAATCAAAATTAACAAGCCAGCAGGGT 783
QY      213 GlnLys-----LysArgSerSerCysCys 220
Db      784 TCCGAGGACCAAGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 819
```

Search completed: June 23, 2003, 19:50:43  
Job time : 231 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 23, 2003, 19:02:43 ; Search time 2487 Seconds  
(without alignments)  
2609.540 Million cell updates/sec

Title: us-09-817-199b-2

Perfect score: 1150

Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESOKKRSCCSFM 223

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame+\_p2n.model -DEV-xlh  
-Q/cgn2\_1/USPTO.spool/US09817199/runat\_18062003\_145102\_795/app.query.fasta\_1.391  
-DB-GenEmbl -OFMT-fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE-LOCAL  
-OUTFW-pto -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-US09817199.ecgn\_1\_1\_2496.erunat\_18062003\_145102\_795 -NCPU=3  
-NO\_MMALP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

RESULT 1

ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                  |
|------------|-------|-------------|--------|-------|------------------------------|
| 1          | 1150  | 100.0       | 1700   | 9     | BC016615 Homo sapi           |
| 2          | 1150  | 100.0       | 1977   | 9     | AK098068 Homo sapi           |
| 3          | 1145  | 99.6        | 1116   | 6     | AX236082 Sequence            |
| 4          | 1081  | 94.0        | 690    | 10    | AF233582 Mus muscu           |
| 5          | 977   | 85.0        | 576    | 6     | AX236084 Sequence            |
| 6          | 780.5 | 67.9        | 1513   | 9     | BC007681 Homo sapi           |
| 7          | 757   | 65.8        | 1320   | 9     | AB027137 Homo sapi           |
| 8          | 754   | 65.6        | 1098   | 10    | U18771 Rattus norv           |
| 9          | 753   | 65.5        | 573    | 9     | AF498952 Homo sapi           |
| 10         | 731   | 63.6        | 1340   | 6     | AK062279 Sequence            |
| 11         | 710.5 | 61.8        | 3105   | 9     | AK054846 Homo sapi           |
| 12         | 649.5 | 56.5        | 2536   | 3     | AY061826 Drosophil           |
| 13         | 521.5 | 45.3        | 1337   | 10    | BC019990 Mus muscu           |
| 14         | 518.5 | 45.1        | 1074   | 8     | AB024994 Ciccr ari           |
| 15         | 517.5 | 45.0        | 2145   | 3     | DDISAS1A M34456 D.discoideu  |
| 16         | 517.5 | 45.0        | 10034  | 2     | AC115606 Dictyoste           |
| 17         | 516   | 44.9        | 638    | 10    | S53270 MEL-RAS-rel           |
| 18         | 516   | 44.9        | 765    | 5     | DYGORA2 M38391 Discopoge o   |
| 19         | 516   | 44.9        | 911    | 8     | AY084345 Arabidops           |
| 20         | 516   | 44.9        | 2048   | 9     | BC002977 Homo sapi           |
| 21         | 516   | 44.9        | 2818   | 9     | AK025165 Homo sapi           |
| 22         | 513.5 | 44.7        | 624    | 9     | AF498943 Homo sapi           |
| 23         | 513.5 | 44.7        | 660    | 9     | HSMRAB8 X56741 H.sapiens m   |
| 24         | 513.5 | 44.7        | 740    | 10    | RU53475 U53475 Rattus norv   |
| 25         | 513.5 | 44.7        | 760    | 4     | CFRAB8 X56385 Canine rab8    |
| 26         | 513.5 | 44.7        | 782    | 8     | PEACRBP09 D13548 Pea mRNA fo |
| 27         | 513.5 | 44.7        | 1265   | 9     | AB038995 Homo sapi           |
| 28         | 512.5 | 44.6        | 1080   | 8     | BVRAB1 Z49152 B.vulgaris     |
| 29         | 512   | 44.5        | 1980   | 9     | S53268 S53268 Homo sapien    |
| 30         | 511.5 | 44.5        | 2497   | 6     | AX285074 Sequence            |
| 31         | 511.5 | 44.5        | 2497   | 6     | AX285080 Sequence            |
| 32         | 511.5 | 44.5        | 2497   | 6     | AX285089 Sequence            |
| 33         | 511.5 | 44.5        | 2497   | 6     | AX285091 Sequence            |
| 34         | 510   | 44.3        | 919    | 8     | DCJ001367 Daucus ca          |
| 35         | 509.5 | 44.3        | 1082   | 8     | LJRAB8D Z73947 L.japonicus   |
| 36         | 509   | 44.3        | 955    | 8     | S66160 S66160 ricl-ras-re    |
| 37         | 509   | 44.3        | 1163   | 8     | LJRAB8A Z73944 L.japonicus   |
| 38         | 508.5 | 44.2        | 1613   | 3     | AB006189 Drosophil           |
| 39         | 508.5 | 44.2        | 2349   | 3     | AY060425 Drosophil           |
| 40         | 507.5 | 44.1        | 979    | 8     | LJRAB1C Z73932 L.japonicus   |
| 41         | 507   | 44.1        | 881    | 8     | AY056351 Arabidops           |
| 42         | 507   | 44.1        | 821    | 8     | AY052333 Arabidops           |
| 43         | 507   | 44.1        | 986    | 8     | AB079023 Nicotiana           |
| 44         | 507   | 44.1        | 1025   | 8     | AY086756 Arabidops           |
| 45         | 507   | 44.1        | 1033   | 8     | AB079021 Nicotiana           |





**TITLE** Direct Submission  
**JOURNAL** Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatar, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
**COMMENT** NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.  
**FEATURES** Location/Qualifiers  
 source  
 1. .1977  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="TRACH2000540"  
 /tissue\_type="trachea"  
 /clone\_lib="TRACH2"  
 /note="cloning vector: pME18SFL3"  
 57. .728  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAC05227.1"  
 /db\_xref="GI:21758002"  
 /translation="MTGPGAVATRGAEAPERSPPSPSYDLTGKVMLLGDTGVGKTC FLQFDGAFSLGTFATGVDFRNKVVTVGVKVLQIWDTAGQERFSRVTHAYRD AQALLLIDITNKSFDNIIRAWLIEHEYAQRDVIIMLGNKADMSSEIRSDGFT LAREYGVPPLETSKTMNVELAFALAKELKYRAGHQADEPSFQIRDYVESQKRSS CCSEFM"  
 BASE COUNT 442 a 560 c 539 g 436 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.16e-108 Length: 1977  
 Score: 1150.00 Matches: 223  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-817-199B-2 (1-223) x AK098068 (1-1977)  
 Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 Db 57 ATGACGGGCGACCGGCGCGGCGTGTCCACCGCGGATGGCGAGCGCCCGGCTCCCG 116  
 Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 Db 117 CCTGCACTCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCGGGAGACACAGGCGTC 176  
 Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 Db 177 GGCAAAACATGTTTCCCTGATCCATTCACAGCGGCGCTTCCTGTCGGGAACCTTCATA 236  
 Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
 Db 237 GCCACCTCGCATAGACTTCAGGAACAAGGTGCTGCTGATGCGGTGAGAGTGAAG 296  
 Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
 Db 297 CTGCAGATCTGGACACCGCTGGCGAGACGGTTCGAGCGGTACCCATGCTTATATAC 356  
 Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
 Db 357 AGAGATGCTCAGGCGCTTCTGCTGATGACATCACCAACAAATCTCTTTCGACAAC 416  
 Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140  
 Db 417 ATCAGGCGCTGGCTCACTGAGATTCAGTATGCCAGAGGCGGTGGTATCATGCTG 476  
 Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160

Db 477 CTAGGCAACAGGCGGATATGAGCAGCGCAAGAGTAGTATCCCTCCGAGACGGAGAGACC 536  
 Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 Db 537 TTGSCCAGGAGTAGTACGGTGTTCCTCTCTGGAGACCGCAGAGCTGGCATGAATGTG 596  
 Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 Db 597 GAGTAGCTTCTTGGCCATCGCCAAAGGAACATAACCGGCGGCGCATCAGCGGAT 656  
 Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerCysCys 220  
 Db 657 GAGCCAGCTTCAGATCCGAGACTATGTAGTCCCAAGAAAGAGCGCTCCAGCTGCTGC 716  
 Qy 221 SerPheMet 223  
 Db 717 TCCATCATG 725  
 RESULT 3  
 AX236082  
 LOCUS AX236082 1116 bp DNA linear PAT 26-SEP-2001  
 DEFINITION Sequence 13 from Patent WO0164887.  
 ACCESSION AX236082  
 VERSION AX236082.1 GI:15795889  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1116)  
 AUTHORS Meyers, R.A.  
 TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins  
 JOURNAL Patent: WO 0164887-A 13 07-SEP-2001;  
 Millennium Pharmaceuticals, Inc. (US)  
 FEATURES Location/Qualifiers  
 source 1. .1116  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 124. .699  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC88255.1"  
 /db\_xref="GI:15795890"  
 /translation="MLLGDVGKTCFLIQKDGAFSLGTFIATVGIDFRNKKVTVDG VRVKLQIWDTAGQERFSRVTHAYRDAQALLLIDITNKSFDNIIRAWLIEHEYAQR DVVIMLGNKADMSSEIRSDGFTLAREYGVPPLETSKTMNVELAFALAKEL YRAGHQADEPSFQIRDYVESQKRSSCCSEFM"  
 BASE COUNT 260 a 313 c 322 g 221 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.35e-108 Length: 1116  
 Score: 1145.00 Matches: 222  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.57% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-817-199B-2 (1-223) x AX236082 (1-1116)  
 Qy 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 21  
 Db 31 ACGGCGACCGCGCGCGCGTTCGCCACCGGATGGCGAGCGCCCGCGCTCCCGGCC 90  
 Qy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41  
 Db 91 TGCATCGGACTAGACCTCAGCGGCAAGGTGATGCTCTCTGGGAGACACAGGCGTCGCG 150  
 Qy 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61  
 Db 151 AAAACATGTTTCCCTGATCCAAATCAAGACGGGCGCTTCTCTCGGGAACCTTCATGCC 210  
 Qy 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81

```

Db      211  ACCGTCGCATAGACTTCAGGACACAGCTGCTGGATGGCTGAGAGTGAAGCTG 270
QY      82  GlnIleTrrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db      271  CAGATCTGGACACCGCTGGGCGAGGACGGTCCGAAGCGTCACCCATGCTTATTACAGA 330
QY      102  AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db      331  GATGCTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 390
QY      122  ArgAlaTrrpLeuThrGluIleHisGlnTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db      391  AGGGCTGGCTCAGTACGATTCATGAGTATGATGATGATGATGATGATGATGATGATGAT 450
QY      142  GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db      451  GGCACAAAGGGGATATGAGCAGCGCAAGAGTGTATCTCTCCGAAGACGGAGACCTTG 510
QY      162  AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db      511  GCCAGGGAGTACGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 570
QY      182  LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db      571  TTAGCCTTCTGGCCATCGCCAAAGAACTGAAATACCGCGCGGCATCAGCGGATCAG 630
QY      202  ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
Db      631  CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCTGCTGCC 690
QY      222  PheMet 223
Db      691  TTCATG 696

RESULT 4
AF233582
LOCUS      690 bp mRNA linear ROD 02-MAY-2000
DEFINITION Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
ACCESSION AF233582
VERSION    AF233582.1 GI:7677421
KEYWORDS   Mus musculus.
SOURCE     Mus musculus.
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 690)
AUTHORS   Masuda,E.S., Luo,Y., Young,C., Shen,M., Rossi,A.B., Huang,B.C.,
          Yu,S., Bennett,M.K., Payan,D.G. and Scheller,R.H.
TITLE     Rab37 is a novel mast cell specific GTPase localized to secretory
          granules
JOURNAL   FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE   20189834
PUBMED   10722846
REFERENCE  2 (bases 1 to 690)
AUTHORS   Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
TITLE     Direct Submission
JOURNAL   Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
          Avenue, South San Francisco, CA 94080, USA
FEATURES   source
            1..690
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /gene="Rab37"
            /gene="Rab37"
            /codon_start=1
            /product="GTPase Rab37"
            /protein_id="AAF67162.1"
            /db_xref="GI:7677422"
            /translation="MTGTPGAATAGDGEAPERSPPFNLDLTKVLLGDSGVGKTC
            FLIQKDAFLSGTFIATVGIDFRKNVTVIDGARVKLIQWDTAGQERFSVTHAYRD

```

```

          AQAALLLYDITNOSSEFNIRAWLTIHEYAQRDQVIVMLLGNKADYSSSERVIRSEDEGT
          LAREYGVPMETSAKTGMNVELAFIALAIKELIRAGQPDPEFPOIRDIYVESQKRSS
          CCSFV*
BASE COUNT 160 a 181 c 201 g 148 t
ORIGIN
Alignment Scores:
Pred. No.: 6,63e-102 Length: 690
Score: 1081.00 Matches: 209
Percent Similarity: 96.41% Conservative: 6
Best Local Similarity: 93.72% Mismatches: 8
Query Match: 94.00% Indels: 0
DB: 10 Gaps: 0
US-09-817-199b-2 (1-223) x AF233582 (1-690)
QY      1  MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db      1  ATGACTGGCACACACGAGGAGCTGTACCCCTGGGGATGGCGAGGCCCTTGAGCGCTCCCGG 60
QY      21  ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db      61  CCTTCAGCCGCGAATACGATCTCACCGGCAAGGTGATGCTCTCTTGAGAGACTCGGGCGTC 120
QY      41  GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db      121  GCAAAACCTGTTCTCTGATCCAAATCAAGACGGGGCTCTCTCTCGGAACTTCATATA 180
QY      61  AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db      181  GCCACCGTCGCATAGACTTCAGGAATAAAGTGTGACAGTGGATGGTCCAGGGTGAGG 240
QY      81  LeuGlnIleTrrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db      241  CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGACGTGTGACCCATGCTTATTAC 300
QY      101  ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db      301  CGAGATGCTCAGGCTTCTCTCTGTTGATGACATCACCAACCCAGCTCTCTTTTGACAAAC 360
QY      121  IleArgAlaTrrpLeuThrGluIleHisGlnTyrAlaGlnArgAspValIleMetLeu 140
Db      361  ATCAGGGCTGGCTGCACAGAGATTCATCAGTATGCCAGAGGGGAGCTGGTATGATGCTT 420
QY      141  LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db      421  CTAGCAACAAGGCGCGATGTAAGCAGCAAGAGGGTGTATCCCTCTCTGAAGATGGAGACA 480
QY      161  LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db      481  CTGGCCAGGAATATGTTGTTCTCTTCATGGAGACCATGTCACCAAGCTGGCATGAACGTG 540
QY      181  GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db      541  GAGTTGGCTTCTTGCAATTCGCAAGAACTGAAATACCGTGCAGGGGAGGAGCGCTGAT 600
QY      201  GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db      601  GAGCCACGCTTCCAGATCCGAGACTATGTGGAGTCCAGAGAGCGCTCCAGAGAGCGCTGC 660
QY      221  SerPheMet 223
Db      661  TCCTTTGTG 669
RESULT 5
AX236084
LOCUS      AX236084
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION    AX236084.1 GI:15795891
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens

```



```

QY 31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
Db 181 AAGGTCATGCTGCTGGGGAGCTCGGGTGTGGGAAGAGCTGCTCTGCTGGATTCAAG 240
QY 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
Db 241 GATGTGCTTCTCTGGCGGGAGCTTCATCTCCACCGTAGGATGACTTCGGGAACAA 300
QY 71 ValValThrValAspGlyValArgValLysLeuGlnIleTTPAspThrAlaGlyGlnGlu 90
Db 301 GTCTCTGGAGCTGATGCTGTAAGTGAAGTGCAGATGTGGACACAGCTGCTCAGGAG 360
QY 91 ArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyr 110
Db 361 CGTTCCGAGGTGTACCATGCTCTACTACCGGATGCTCATGCTCTGCTGCTCTAC 420
QY 111 AspileThrAsnLysSerSerPheAspAsnIleArgAlaThrLeuThrGluIleHisGlu 130
Db 421 GATGTCCACCAACAGGCTCTCTTGACACATCCAGGCTGGCTGACCGAGATCCAGAG 480
QY 131 TyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
Db 481 TACGCCACGACGAGCTGCTGCTCATGCTGTGGGAACAAAGTGGACTCTGCCATGAG 540
QY 151 ArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeu 170
Db 541 CGTGTGGTAGAGGAGGAGCGGGAGAGCTGGCCAAAGGAGTATGGACTGCCCTTCATG 600
QY 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
Db 601 GAGACCAGCGCAAGACGGGCTCAACGTGGACTTGGCTTCACAGCCATAGCAAGAGAG 660
QY 191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
Db 661 TTGAAGCAGGCTCCATGAAGGCTCCACGAGCGCGCTCCGCTGCATGATACGTT 720
QY 211 GluSerGlnLysLysArgSerSerCysCys 220
Db 721 AAGAGGAGGTCGAGGCGCTCTCTGCTGC 750

RESULT 7
AB027137
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
gene
CDS

AB027137 1320 bp mRNA linear PRI 29-SEP-1999
Homo sapiens v46133 mRNA for RAB-26, complete cds.
AB027137.1 GI:5931611
RAB-26; v46133.
Homo sapiens cDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1320)
Miyajima, N., Seki, N., Hattori, A., Hayashi, A., Kozuma, S.,
Muramatsu, M. and Saito, T.
Human RAS-related protein RAB-26
Published only in Database (1999)
2 (bases 1 to 1320)
Miyajima, N., Seki, N., Hattori, A., Hayashi, A., Kozuma, S.,
Muramatsu, M. and Saito, T.
Direct Submission
Submitted (11-MAY-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group; Inage-ku Anagawa
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp,
Tel: 81-43-201-3135, Fax: 81-43-251-9818)
Location/Qualifiers
1. .1320
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .1320
/feature="v46133"
4. .576
/feature="v46133"

```

```

/note="RAS-related protein"
/codon_start=1
/product="RAB-26"
/protein_id="BAA84707.1"
/db_xref="GI:5931612"
/translation="MLVDSGVKTCILVRFKDFAGLFTISTVIGIDFRNKVLVDVG
VKVQLQMDTATGQERFSTHAYYHDAHALLLYDVTNKASFNDFOALWTEHYEAOH
DVALLMLGNKVDYSAHERVVKREDGEKLAKEYGLPMPETSAKTGLNVDLAFATAIKELK
QRSMKAPSEPRFLHDYVKRGRGASCCRP"
BASE COUNT 273 a 379 c 395 g 273 t
ORIGIN

Alignment Scores:
2.78e-68 Length: 1320
Score: 757.00 Matches: 139
Percent Similarity: 88.36% Conservative: 28
Best Local Similarity: 73.54% Mismatches: 22
Query Match: 65.83% Indels: 0
DB: 9 Gaps: 0

US-09-817-199B-2 (1-223) x AB027137 (1-1320)
QY 32 ValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysasp 51
Db 1 GTCATGCTGTGGGGAGCTCGGGTGTGGGAAGAGCTGCTGCTGCTGCTCAAGGAT 60
QY 52 GlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysVal 71
Db 61 GGTGCTTCTCTGGCGGAGCTTCATCTCCACCGTAGGATGACTTCCGGAACAAAGTT 120
QY 72 ValThrValAspGlyValArgValLysLeuGlnIleTTPAspThrAlaGlyGlnGluArg 91
Db 121 CTGGACGTGGATGGTGTGAAGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGAGCGG 180
QY 92 PheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAsp 111
Db 181 TTCGCAGCTGTACCATGCTCTACTACCGGATGCTCATGCTGCTGCTGCTTACCAGT 240
QY 112 IleThrAsnLysSerSerPheAspAsnIleArgAlaThrLeuThrGluIleHisGluTyr 131
Db 241 GTCACCAACAGGCTCTCTTGACAAACATCCAGGCTGGCTGACCGAGATCCACAGATAC 300
QY 132 AlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArg 151
Db 301 GCCCAGCAGCAGCTGGCTCATGCTGTGGGAACAAAGTGGACTCTGCCATGAGCGT 360
QY 152 ValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGlu 171
Db 361 GTGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 172 ThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeu 191
Db 421 ACCAGCGCAAGAGCGGCTCAACGTGGACTTGGCTTTCACAGCCATACCAAGAGGTTG 480
QY 192 LysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGlu 211
Db 481 AAGCAGCGCTCCATGAAGGCTCCAGCGAGCGGCGGCTCCGCTGCTGCTGCTTACGTTAG 540
QY 212 SerGlnLysLysArgSerSerCysCys 220
Db 541 AGGAGGAGGTCGAGGCGCTCTCTGCTGC 567

RESULT 8
AB027137
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
gene
CDS

AB027137 1098 bp mRNA linear ROD 18-JUL-1995
Rattus norvegicus Rab26 mRNA, complete cds.
AB027137.1 GI:619733
Rattus norvegicus
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```

```

REFERENCE 1 (bases 1 to 1098)
AUTHORS Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
TITLE Molecular cloning of a new member of the Rab protein family, Rab
26, from rat pancreas
JOURNAL Biochem. Biophys. Res. Commun. 207 (3), 950-956 (1995)
MEDLINE 95169156
PUBMED 7864900
REFERENCE 2 (bases 1 to 1098)
AUTHORS Williams,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of
Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA
FEATURES
source
1..1098
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="pancreas"
/dev_stage="adult"
1..29
30..602
/codon_start=1
/product="Rab26"
/protein_id="AAA69955.1"
/db_xref="GI:619734"
/translation="MLVDSGVGKTCILLVRFKDGAGLAGTFFISTVGIDFRNKVLVDVG
MKVKLQIWDTAGQERFRSVTHAYTRDAHALLLDITNKDSFDNIQAWLFEIQAYQ
DVLMLLGNKVDSPQERVKREDGEKLAKEVGLPFMTSAKSLGNLVDLAFATIAKELK
ORSTKAPSEPRFLHDYVKREGVSCRL"
603..1098
3'UTR 1063..1098
polyA_signal 267 a 277 c 306 g 248 t
ORIGIN
1
4,44e-68 Length: 1098
754.00 Matches: 140
87.89% Conservativity: 27
73.68% Mismatches: 23
65.57% Indels: 0
10 Gaps: 0
Alignment Scores:
Pred. No.: 4,44e-68 Length: 1098
Score: 754.00 Matches: 140
Percent Similarity: 87.89% Conservativity: 27
Best Local Similarity: 73.68% Mismatches: 23
Query Match: 65.57% Indels: 0
DB: 10 Gaps: 0
US-09-817-199B-2 (1-223) x RNU18771 (1-1098)
Qy 31 LysValMetLeuLeuLeuAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
|||||
Db 24 AAGTCATGCTGGTGGGGATTCCGGTGTGGGAAGACGTCCTGCTCGTCCGCTCAAG 83
Qy 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
|||||
Db 84 GATGGGCGCTTCCTGGTGGTACCTTCATCTCCACTGTGGGCATCGACTTCGGAAATAA 143
Qy 71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
|||||
Db 144 GTTCTGGATGGATGGATGAAGGTGAAGCTGCAGATCTGGGACCGCTGGTCAGGAG 203
Qy 91 ArgPheArgSerValThrHisAlaThrTyrArgAspAlaGlnAlaLeuLeuLeuTyr 110
|||||
Db 204 CGGTTCGGAAGTGCACCATGCTACTACCGGATGCTACCGCACCTGTTGCTGCTCTAC 263
Qy 111 AspileThrAsnLysSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
|||||
Db 264 GACATCAACCAAGAATTCCTTCGACAACATCAGCGCTGGTGGACAGAAATCCAGGA 323
Qy 131 TyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
|||||
Db 324 TATCCCAAGCAGGACGTGGTCTCTGCTGTGGGACACAGGTGTGACTCTACTCAAGAA 383
Qy 151 ArgValIleArgSerGlyAspGlyGluThrThrAlaArgGluTyrGlyValProPheLeu 170
|||||
Db 384 CGATGTTAAGAGGAGATGGGAGAAATACCAAGGAGTAGTGGCTCCCATCATG 443
Rattus
1 (bases 1 to 1098)
AUTHORS Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
TITLE Molecular cloning of a new member of the Rab protein family, Rab
26, from rat pancreas
JOURNAL Biochem. Biophys. Res. Commun. 207 (3), 950-956 (1995)
MEDLINE 95169156
PUBMED 7864900
REFERENCE 2 (bases 1 to 1098)
AUTHORS Williams,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of
Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA
FEATURES
source
1..1098
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="pancreas"
/dev_stage="adult"
1..29
30..602
/codon_start=1
/product="Rab26"
/protein_id="AAA69955.1"
/db_xref="GI:619734"
/translation="MLVDSGVGKTCILLVRFKDGAGLAGTFFISTVGIDFRNKVLVDVG
MKVKLQIWDTAGQERFRSVTHAYTRDAHALLLDITNKDSFDNIQAWLFEIQAYQ
DVLMLLGNKVDSPQERVKREDGEKLAKEVGLPFMTSAKSLGNLVDLAFATIAKELK
ORSTKAPSEPRFLHDYVKREGVSCRL"
603..1098
3'UTR 1063..1098
polyA_signal 267 a 277 c 306 g 248 t
ORIGIN
1
4,44e-68 Length: 1098
754.00 Matches: 140
87.89% Conservativity: 27
73.68% Mismatches: 23
65.57% Indels: 0
10 Gaps: 0
Alignment Scores:
Pred. No.: 4,44e-68 Length: 1098
Score: 754.00 Matches: 140
Percent Similarity: 87.89% Conservativity: 27
Best Local Similarity: 73.68% Mismatches: 23
Query Match: 65.57% Indels: 0
DB: 10 Gaps: 0
US-09-817-199B-2 (1-223) x AF498952 (1-573)
Qy 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
|||||
Db 1 ATGTTGGTGGGGGACTCGGGTGGGGAAGACCTGCTGCTGGTGGGATTCAAGATGCT 60
Qy 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
|||||
Db 61 GCTTCTCTGGCGGGACCTTCATCTCCACCGTAGGCATTCGCTCCGGAACAAGTCTTG 120
Qy 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
|||||
Db 444 GAGACCAAGCGCAAGAGCGGCTCAATGTGACTTGGCTTTTACAGCCATAGCAAGAGG 503
Qy 191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
|||||
Db 504 CTGAACAAGATCCACCAAGCTCCAGTCCGCGCTTCAGGCTCATGACTATGTG 563
Qy 211 GluSerGlnLysLysArgSerSerCysCys 220
|||||
Db 564 AAGAGGAGGAGCCGAGGGGTCTCTGCTGT 593
RESULT 9
AF498952
LOCUS Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete cds.
DEFINITION AF498952.1 GI:20379079
ACCESSION AF498952
VERSION AF498952.1
KEYWORDS Homo sapiens
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Homo sapiens RAB family small GTP binding protein RAB26
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES
Location/Qualifiers
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
1..573
/gene="RAB26"
1..573
/gene="RAB26"
/feature="RAB family member"
/codon_start=1
/product="small GTP binding protein RAB26"
/protein_id="AA21100.1"
/db_xref="GI:20379080"
/translation="MLVDSGVGKTCILLVRFKDGAGLAGTFFISTVGIDFRNKVLVDVG
VKVLQIWDTAGQERFRSVTHAYTRDAHALLLDITNKDSFDNIQAWLFEIHTAQH
DVALMLLGNKVDSPQERVKREDGEKLAKEVGLPFMTSAKSLGNLVDLAFATIAKELK
ORSMKAPSEPRFLHDYVKREGVSCRP"
BASE COUNT 117 a 151 c 191 g 114 t
ORIGIN
1
2,41e-68 Length: 573
753.00 Matches: 138
88.30% Conservativity: 28
73.40% Mismatches: 22
65.48% Indels: 0
9 Gaps: 0
Alignment Scores:
Pred. No.: 2,41e-68 Length: 573
Score: 753.00 Matches: 138
Percent Similarity: 88.30% Conservativity: 28
Best Local Similarity: 73.40% Mismatches: 22
Query Match: 65.48% Indels: 0
DB: 9 Gaps: 0
US-09-817-199B-2 (1-223) x AF498952 (1-573)
Qy 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
|||||
Db 1 ATGTTGGTGGGGGACTCGGGTGGGGAAGACCTGCTGCTGGTGGGATTCAAGATGCT 60
Qy 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
|||||
Db 61 GCTTCTCTGGCGGGACCTTCATCTCCACCGTAGGCATTCGCTCCGGAACAAGTCTTG 120

```

QY 73 ThrValAspGlyValArgValLysLeuGlnIleThrPaspThrAlaGlyGlnGluArgPhe 92  
 DB 121 GACGTGGATGGTGAAGCTGAAGTGGACATGCTGGGACACAGCTGGTCAGGAGCGGTC 180  
 QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112  
 DB 181 CCGAGTGTACCATGCTTACCTACCGGATGCTCATGCTGCTGCTCATGCTGCTCATGCTGCT 240  
 QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTyrLeuThrGluIleHisGluTyrAla 132  
 DB 241 ACCAACAGGCGCTCTTTGACACATCCAGGCTGGTGCACCGAGATCCACGAGTACGCC 300  
 QY 133 GlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152  
 DB 301 CAGCAGCAGCGGCTCATGCTGCTGGGAAACAAGGTGGACTCTGCCATGAGGCTGTG 360  
 QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172  
 DB 361 GTGAAGAGGAGGAGCGGAGAGCTGGCCAGGAGTATGAGTCCCTTCATGGAGACC 420  
 QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaLysGluLeuLys 192  
 DB 421 AGCGCCAAAGCGGCTCAACCTGGACTTGGCTTCACAGCCATAGCAAGGAGTTGAAG 480  
 QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212  
 DB 481 CAGCGCTCATGAGGCTCCAGGAGCGCGCTTCCGCTCATGATACCTTAAAGAGG 540  
 QY 213 GlnLysLysArgSerSerCysCys 220  
 DB 541 GAGGCTCGAGGCGCTCTCTGCTGC 564

RESULT 10  
 AR062279  
 LOCUS AR062279 1340 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 2 from patent US 5843717.  
 ACCESSION AR062279  
 VERSION AR062279.1 GI:5989970  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1340)  
 AUTHORS Hillman,J.L. and Guegler,K.J.  
 TITLE Rab protein  
 JOURNAL Patent: US 5843717-A 2 01-DEC-1998;  
 FEATURES  
 source Location/Qualifiers  
 1. 1340  
 /organism="unknown"  
 BASE COUNT 276 a 387 c 403 g 274 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.32e-65 Length: 1340  
 Score: 731.00 Matches: 141  
 Percent Similarity: 86.73% Conservative: 29  
 Best Local Similarity: 71.94% Mismatches: 25  
 Query Match: 63.57% Indels: 2  
 DB: 6 Gaps: 0

US-09-817-1999b-2 (1-223) x AR062279 (1-1340)

QY 26 TyrAspLeuThrGlyLysValMetLeuGlyAspThrGlyValGlyLysThrCysPhe 45  
 DB 22 TAGAGCTGCCTTCAAGGATGATGCTGGTGGGGACTCGGGTGGGGAAGACCTGTCTG 81  
 QY 46 Leu-IleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIle 65  
 DB 82 CTGGTGGCATCAAGGATGGTGTCTTCTGGGGGACCTTCATCTCCACCGTAGC-AT 140  
 QY 65 eAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleThrPhe 85  
 DB 141 TGACTCCCGGAACAAAGTCTCGAGCTGTGAGAGTGAAGCTGCAGATGTGGGA 200

QY 85 pThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAl 105  
 DB 201 CACAGCTGCTCAGGAGCGGTTCCGAGTGTACCATGCTACTACCGGATGCTCATGC 260  
 QY 105 aleuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpIle 125  
 DB 261 TCTGCTGCTGCTCAGGATGTCCACAAAGGCTCTTTCACAAACATCCAGGCTGTGCT 320  
 QY 125 uThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAl 145  
 DB 321 GACCGAGATCCAGGATGACGCCACGACGACGCTGGCGCTCATGCTGCTGGGGAACAAGT 380  
 QY 145 aAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaPheLe 165  
 DB 381 GGACTCTGCCATGAGCGTGTGTGAAGAGGAGGAGCGGAGAGCTGGCCAAAGGAGTA 440  
 QY 165 rGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLe 185  
 DB 441 TGGACTGCTCTCATGAGAGACCGCCCAAGCGGCTCAACGTGGACTTGGCGCTTCAC 500  
 QY 185 uAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheG 205  
 DB 501 AGCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGCTCCACGAGCGCGCTTCGG 560  
 QY 205 nIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
 DB 561 GCTGCTCATGATTACGTTAAGAGGAGGCTCGAGGGGCTCTCTGCTGC 606

RESULT 11  
 AK054846

LOCUS AK054846 3105 bp mRNA linear PRI 01-AUG-2002  
 DEFINITION Homo sapiens cDNA FLJ30284 f1s, clone BRACE2002812, moderately similar to Mus musculus GTPase Rab37 (Rab37) mRNA.

ACCESSION AK054846

VERSION AK054846.1 GI:16549463

KEYWORDS oligo capping; f1s (full insert sequence).

SOURCE Homo sapiens cerebellum cDNA to mRNA, clone\_lib:BRACE2

clone:BRACE2002812.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,

Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,

Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,

Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,

Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,

Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,

Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.

NEDO human cDNA sequencing project

Unpublished

REFERENCE 2 (bases 1 to 3105)

Isogai,T., Otsuki,T. and Sugiyama,T.

Direct Submission

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: RAB and

HRI.

FEATURES

Location/Qualifiers

1. 3105

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="BRACE2002812"

/tissue\_type="cerebellum"

/clone\_lib="BRACE2"

/notes="cloning vector: pME18SFL3"

|  |        |  |       |       |
|--|--------|--|-------|-------|
| BASE COUNT                                   | 680 a  | 905 c  | 898 g | 622 t |
| ORIGIN                                       |        |  |       |       |
| Alignment Scores:                            |        |  |       |       |
| Pred. No.:                                   | 5e-63  | Length:  | 3105  |       |
| Score:                                       | 710.50 | Matches:   | 182   |       |
| Percent Similarity:                          | 39.47% | Conservative:  | 13    |       |
| Best Local Similarity:                       | 36.84% | Mismatches:  | 26    |       |
| Query Match:                                 | 61.78% | Indels:  | 274   |       |
| DB:  | 9      | Gaps:  | 5     |       |
| US-09-817-199b-2 (1-223) x AK054846 (1-3105) |        |  |       |       |
| Qy   | 3      | GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- | 21    |       |
| Db   | 375    | GGGAGAGCCGGAGCGGAGCTCAAGCCCTGGCATGACCTGACGACCGGATTCCTAC      | 434   |       |
| Qy   | 22     | -----CysSerProSerTyrAsp-----LeuThrGlyLysValMetLeuLeuGly      | 36    |       |
| Db   | 435    | CAGGAGGAGCTGGCCCTGACTCAAGCAGCCAGCTGCTGATAGACCATCTCTGGTGGGT   | 494   |       |
| Qy   | 37     | AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer | 56    |       |
| Db   | 495    | GACAGTGGTGGGAAGACGCTCTCTGCTGGTTCAGTTCGATCAGGCGCAAGTTCATCCCC  | 554   |       |
| Qy   | 57     | GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGly | 76    |       |
| Db   | 555    | GGCTCTCTTCGGCCACTGCGGCATCGGATTCAGCAACAAGGTGGTGAATGGATGGC     | 614   |       |
| Qy   | 77     | ValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThr | 96    |       |
| Db   | 615    | GTGAGAGTGAAGCTGAGATCTGGGACACCGCTGGGCGAGGAGCGGTTCGAAGCGTCACC  | 674   |       |
| Qy   | 97     | HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer | 116   |       |
| Db   | 675    | CATGCTTATTACAGAGATGCTCAGGCGCTCTCTCTGCTGATGACATCACCACAAATCT   | 734   |       |
| Qy   | 117    | SerPheAsnIleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspVal    | 136   |       |
| Db   | 735    | TCTTTCGACAACTACAGGCGCTGGCTCAGTTCATGATTCATGATGATGCGCCAGAGCGTG | 794   |       |
| Qy   | 137    | ValIleMetLeuLeuGlyAsnLys-----                                | 144   |       |
| Db   | 795    | GTGATCATGTCTGTAGCAACAAGGTGAGTGGCTCCGGGCGCAGGTCAGCCAGCCCTGC   | 854   |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 855    | ACTTCTCAGCCCTAGCGGCCCCCAATAACCAACCAACAGATTATCTAGGATCTCTCC    | 914   |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 915    | TGAAAGGAGCTCTGCAGCCCTCCAGCTCAGGCGGTGAGATATCTGGAGGCTTCTGCCCAT | 974   |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 975    | CCCATGTGCCCTTCCAGGGAAGTCCAAAGTTGTGCTGAGAAATCAAGGGTGGCCAG     | 1034  |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 1035   | TTCTCAGCCCCATTAGACGAGTGNACAGGTCAGGTCAGGCGCTAAGAGTGCAAA       | 1094  |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 1095   | GGGTTAGCCCCAACTGCTGTCTTATTCAGAGCCCTTTACCAAAAGGTGAGATCCAGAGC  | 1154  |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 1155   | TGGAGGTACACTGGGCGAGAAACCTGGCCCCAGGCCAATCACACCTGCTGAGTCCCT    | 1214  |       |
| Qy   | 144    | -----  | 144   |       |
| Db   | 1215   | TGGGCCACAGAGGGCAGCAACCGCTGCTTCTGGGCAAAATATGGCCCGCTGGG        | 1274  |       |

|            |  |   |                        |        |
|------------|--|---|------------------------|--------|
| QY         | 145  | -----   | -----AlaAspMetSerSerG1 | 150    |
| Db         | 1275   | GCAGAGGCTTCCCTCCAGAGTGACCCATTGGGCTTGACAGCGGATATGACACCGA       | 1334                   |        |
| QY         | 150  | uAa3ValIleArgSerGluAspGlyGluThrLeuAlaArg-----                 | 163                    |        |
| Db         | 1335   | AAGAGTGATCCGTTCCGAGACAGGAGAGACTTGGCCAGGGTAAGTGATTGTCTGTGGGA   | 1394                   |        |
| QY         | 164  | -----   | -----G1                | 164    |
| Db         | 1395   | CAGSGTGAAGGTGGGGCAACCCGACGCTGGCCCTGAGGACACTCTCTCCCGGGCAGGA    | 1454                   |        |
| QY         | 164  | uTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe | 184                    |        |
| Db         | 1455   | GTACGGTGTTCCTTCTCTGGAGACAGCGCCAGACTGGCATGAATGTGGAGTTAGCCTT    | 1514                   |        |
| QY         | 184  | eLeuAlaIleAlaLys-----   | -----                  | 189    |
| Db         | 1515   | CTGGGCATCGCCAA-GTGAGAGCTGGGAGGAGGAAGTGTGCGGGGCGAGGGCGGCA      | 1573                   |        |
| QY         | 189  | -----   | -----                  | 189    |
| Db         | 1574   | CACCTCCAGGAATCCAGTAGGGCCCGCCCTGGCCAGCCCTGGACACACCTGCATTCT     | 1633                   |        |
| QY         | 189  | -----   | -----                  | 189    |
| Db         | 1634   | GCAGGCTGAGGTCCATTGCTCTGGGAGCACTGGGCGACTGGGAGAGGGGCGGCGC       | 1693                   |        |
| QY         | 190  | -----   | -----G1                | 190    |
| Db         | 1694   | TCAGCTCTCACCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC      | 1753                   |        |
| QY         | 190  | uLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVa  | 210                    |        |
| Db         | 1754   | ACTGAATACCGGCGGCGCATCAGCGGATGAGCCAGCTTCAGATCCGAGACTATGT       | 1813                   |        |
| QY         | 210  | lGluSerGlnLysLysArgSerCysCysSerPheMet                         | 223                    |        |
| Db         | 1814   | AGAGTCCAGAAAGCGCTCCAGCTGCTGCTCTCTCATG                         | 1853                   |        |
| RESULT 12  |  |   |                        |        |
| LOCUS      | AY061826   | 2536 bp   | mRNA                   | linear |
| DEFINITION | Drosophila melanogaster  | GH21984   | full length            | cdNA.  |
| ACCESSION  | AY061826   |   |                        |        |
| VERSION    | AY061826.1   | GI:16902019   |                        |        |
| KEYWORDS   | FLI CDNA.  |   |                        |        |
| SOURCE     | Drosophila melanogaster.   |   |                        |        |
| ORGANISM   | Drosophila melanogaster  |   |                        |        |
| REFERENCE  | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.   |   |                        |        |
| AUTHORS    | Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarini, H., Li, P., Liao, G., Miranda, A., Mungai, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.  |   |                        |        |
| TITLE      | Direct Submission  |   |                        |        |
| JOURNAL    | Submitted (02-NOV-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA   |   |                        |        |
| COMMENT    | Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720   |   |                        |        |
|            | This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cdNA clone. However, there are |   |                        |        |





```

/protein_id="AAH19990.1"
/db_xref="GI:18043409"
/translation="MAKTYDYLFKLLIGDSGVGKTCVLFKFSDFAFNSTFIISTIGID
FKRTIEIDGRKIKLQIWDTAGQERFRTITAYRGAMGLVYIDTNEKSFQNRNW
IRNIEHASADVEKMLGNKCDVNDKROVSKGEKGLALDYGIKFMETSAKANINVEN
AFTLARDIRAKWDKLEGNPSGSSHGVIIVQOQRISFFRCSL"
BASE COUNT 367 a 338 c 341 g 291 t
ORIGIN

Alignment Scores:
Pred. No.: 4.19e-44 Length: 1337
Score: 521.50 Matches: 101
Percent Similarity: 69.76% Conservative: 42
Best Local Similarity: 49.27% Mismatches: 55
Query Match: 45.35% Indels: 7
DB: 10 Gaps: 3

US-09-817-199B-2 (1-223) x BC019990 (1-1337)

Qy 25 SerTyrAspLeuThrGlyLysValMetLeuGlyAspThrGlyValGlyLysThrCys 44
Db 10 ACCTACGATACCTGTTCAAGCTGCTGCTATCGGGGACTCGGGGGTAGGGAACCTGT 69
Qy 45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64
Db 70 GTCTGTTCGCTCTCCGAGGAGCGCTTC---AACTCCACATTCATCTCTACCATAGGA 126
Qy 65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrp 84
Db 127 ATTGACTTTAAATATTAGGACCATAGAGCTCGATGCGAAGAGATTAACCTGCAGATGG 186
Qy 85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrArgAspAlaGln 104
Db 187 GACACGCGCGCCAGGAGCGGTTTCGAACAACTACAGCAGCGCTACTACGGGGTCCATG 246
Qy 105 AlaLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsnIleArgAlaTrp 124
Db 247 GGTATCATGTGCTCTAGACATTCACCAAGAGAGTCCITTTGACACATCCGGAATGG 306
Qy 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLys 144
Db 307 ATTCGGAACATTGAAGAGCATGCTCTGCAGACGTGGAGAGATGATACTGGGAATAAG 366
Qy 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
Db 367 TGTGATGTGAATGACAGAGACAGGTGTCCAAAGCAACGGGGAGAAAAGCTTGCCTCGAC 426
Qy 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
Db 427 TATGGATCAAGTTTCATGGAGACCGAGTGCACCAAGCCCAACATTAATGTGGAGAAATG 486
Qy 185 LeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGlyHisGlnAla 199
Db 487 TTCACCTTCCAGGGATATCAAGCAAAAATGACAAAATAATTTGGAAGGCAAGCCGCG 546
Qy 200 AspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCys 219
Db 547 CAGGGGAGCAGCCATCAGATCAAGTCAAGTACATGAGCAGCAGCAAGAGAGACGCTTCTTC 606
Qy 220 ---CysSerPheMet 223
Db 607 CGGTGCAGTCTCCTG 621

RESULT 14
AB024994
LOCUS
DEFINITION Cicer arietinum mRNA for rab-type small GTP-binding protein,
complete cds.
ACCESSION AB024994
VERSION AB024994.1 GI:4586579
KEYWORDS rab-type small GTP-binding protein.
SOURCE Cicer arietinum (strain:ILC3279) leaf cDNA to mRNA, clone:INR134.
ORGANISM Cicer arietinum

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
Cicer.
1 (sites)
Ichinose,Y., Tiemann,K., Schwenger-Erger,C., Toyoda,K., Hein,F.,
Hanselle,R. and Barz,W.
Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and
Elicited Cell Cultures as Detected by Differential
cDNA-Hybridization
Unpublished
2 (bases 1 to 1074)
Ichinose,Y.
Direct Submission
Submitted (16-MAR-1999) Yuki Ichinose, Okayama University, Faculty
of Agriculture; Tsushima-naka, 1-1-1, Okayama, Okayama 700-8530,
Japan. (E-mail:yuki@cc.okayama-u.ac.jp, Tel:+81-86-251-8308,
Fax:+81-86-251-8308)
FEATURES
Location/Qualifiers
1..1074
/organism="Cicer arietinum"
/strain="ILC3279"
/db_xref="taxon:3827"
/clone="INR134"
/tissue_type="leaf"
107..115
/codon_start=1
/product="rab-type small GTP-binding protein"
/protein_id="BAA76422.1"
/db_xref="GI:4586580"
/translation="MNPEYDYLFLKLLIGDSGVGKSCLLLRFPADDSYLDVSIISTIGVD
FKRTVSDGKTIKLIQIWDTAGQERFRTITSSYIRGAHGLIIVYDVTQDESFNKQW
LNEIDRVASENVKLLGNKCDLAANKVVSSETAKAFADIGIPFMETSAKNATNVEQ
AFMAAAEIKRMAQSOPANNARPPTVQIRGPMQNKSGCCST"
1074
polya_site
/note="8 A nucleotides"
BASE COUNT 310 a 223 c 207 g 334 t
ORIGIN

Alignment Scores:
Pred. No.: 6.4e-44 Length: 1074
Score: 518.50 Matches: 97
Percent Similarity: 69.90% Conservative: 47
Best Local Similarity: 47.09% Mismatches: 55
Query Match: 45.09% Indels: 7
DB: 8 Gaps: 3

US-09-817-199B-2 (1-223) x AB024994 (1-1074)

Qy 20 ProProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39
Db 101 CCCCTCATGAATCCGATATGACATTTTGTTCAGCTTTTGTTCAGATTTGAGATTCGTGT 160
Qy 40 ValGlyLysThrCysPhePheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59
Db 161 GTGGCAAGTCATGCTCTCCTCCTGAGTTTCTGATGATTCATATCTTGAC---AGCTAT 217
Qy 60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79
Db 218 ATCAGTACAAATTGGAGTGAGCTTTTAAATTCGCATGTTGAGCAAGCGGGAAGACCATT 277
Qy 80 LysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99
Db 278 AAACCTCAAAATTTGGGACACTGCTGGTCAAGAACCTTTCCGAGCTATCTACTAGCAGCTAC 337
Qy 100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAsp 119
Db 338 TATCTGGGGCTCATGGCATTAATTTGTTTATGATGTCACCTACCAAGAGAGCTTTAAC 397
Qy 120 AsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMet 139
Db 398 AATTTAGCAGTGGCTGAATGAATTCACCGCTTATGCAAGTGAAGAAATGTAACACAGCTT 457

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:48:33 ; Search time 49 Seconds  
(without alignments)  
492.451 Million cell updates/sec

Title: US-09-817-199b-2

Perfect score: 1150

Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFPM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCRT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCRTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 1150  | 100.0       | 223    | 10 | US-09-817-199a-2   |
| 2          | 1144  | 99.5        | 226    | 9  | US-09-764-868-684  |
| 3          | 1133  | 98.0        | 222    | 9  | US-09-764-868-1106 |
| 4          | 1081  | 94.0        | 223    | 10 | US-09-817-199a-4   |
| 5          | 977   | 85.0        | 191    | 10 | US-09-794-257-14   |
| 6          | 977   | 85.0        | 191    | 12 | US-10-051-986-3    |
| 7          | 547   | 47.6        | 106    | 10 | US-09-867-550-1812 |
| 8          | 535   | 46.5        | 139    | 9  | US-09-764-868-688  |
| 9          | 501   | 43.6        | 207    | 10 | US-09-794-257-8    |
| 10         | 500   | 43.5        | 190    | 10 | US-09-822-860-5    |
| 11         | 498   | 43.3        | 218    | 10 | US-09-925-300-1571 |
| 12         | 497.5 | 43.3        | 162    | 10 | US-09-834-765-766  |
| 13         | 483   | 42.0        | 246    | 10 | US-09-925-302-534  |
| 14         | 472   | 41.0        | 201    | 10 | US-09-867-736-8    |
| 15         | 469.5 | 40.8        | 224    | 9  | US-10-102-806-466  |
| 16         | 468   | 40.7        | 201    | 10 | US-09-967-736-3    |
| 17         | 466   | 40.5        | 198    | 10 | US-09-794-257-16   |
| 18         | 466   | 40.5        | 198    | 10 | US-09-945-173-5    |
| 19         | 466   | 40.5        | 198    | 10 | US-09-972-529-4    |

|    |       |      |     |    |                    |                    |
|----|-------|------|-----|----|--------------------|--------------------|
| 20 | 458.5 | 39.9 | 222 | 10 | US-09-820-003A-4   | Sequence 4, Appl1  |
| 21 | 448.5 | 39.0 | 212 | 10 | US-09-350-874-67   | Sequence 67, Appl1 |
| 22 | 435   | 37.8 | 401 | 9  | US-09-764-868-701  | Sequence 701, App  |
| 23 | 429.5 | 37.3 | 188 | 9  | US-09-764-868-1120 | Sequence 1120, Ap  |
| 24 | 429   | 37.3 | 212 | 10 | US-09-817-198A-2   | Sequence 2, Appl1  |
| 25 | 427.5 | 37.2 | 212 | 10 | US-09-817-198A-4   | Sequence 4, Appl1  |
| 26 | 416   | 36.2 | 218 | 10 | US-09-817-198A-5   | Sequence 5, Appl1  |
| 27 | 413.5 | 36.0 | 307 | 9  | US-09-764-868-1100 | Sequence 1100, Ap  |
| 28 | 413.5 | 36.0 | 312 | 10 | US-09-925-302-783  | Sequence 2, Appl1  |
| 29 | 386   | 33.6 | 832 | 10 | US-09-834-765-2    | Sequence 1112, Ap  |
| 30 | 377.5 | 32.8 | 222 | 9  | US-09-764-868-1112 | Sequence 692, App  |
| 31 | 377.5 | 32.8 | 225 | 9  | US-09-764-868-692  | Sequence 5, Appl1  |
| 32 | 373.5 | 32.5 | 213 | 10 | US-09-794-257-5    | Sequence 45, Appl1 |
| 33 | 365   | 31.7 | 208 | 9  | US-10-108-605-45   | Sequence 5, Appl1  |
| 34 | 364.5 | 31.7 | 624 | 10 | US-09-834-765-5    | Sequence 762, App  |
| 35 | 364.5 | 31.7 | 625 | 10 | US-09-834-765-762  | Sequence 765, App  |
| 36 | 360.5 | 31.3 | 168 | 10 | US-09-834-765-765  | Sequence 64, Appl1 |
| 37 | 360   | 31.3 | 213 | 9  | US-10-036-542-64   | Sequence 89, Appl1 |
| 38 | 360   | 31.3 | 217 | 9  | US-10-036-542-89   | Sequence 1077, Ap  |
| 39 | 360   | 31.3 | 239 | 10 | US-09-925-301-1077 | Sequence 3, Appl1  |
| 40 | 359   | 31.2 | 217 | 10 | US-09-988-974-3    | Sequence 8, Appl1  |
| 41 | 358.5 | 31.2 | 213 | 10 | US-09-988-974-8    | Sequence 1364, Ap  |
| 42 | 352.5 | 30.7 | 217 | 10 | US-09-925-300-1364 | Sequence 2, Appl1  |
| 43 | 350   | 30.4 | 201 | 10 | US-09-822-860-2    | Sequence 763, App  |
| 44 | 349.5 | 30.4 | 161 | 10 | US-09-834-765-763  | Sequence 6, Appl1  |
| 45 | 344.5 | 30.0 | 211 | 12 | US-10-051-986-6    |                    |

#### ALIGNMENTS

##### RESULT 1

US-09-817-199a-2

; Sequence 2, Application US/09817199A

; Patent No. US20020142380A1

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01187

; CURRENT APPLICATION NUMBER: US/09/817,199A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Human

US-09-817-199a-2

Query Match 100.0%; Score 1150; DB 10; Length 223;  
Best Local Similarity 100.0%; Pred. No. 2e+108;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | MTGTPGAVATRDGEAPERSPPCSYDLTGKVMILGDTGVGKTCFLIOFKDGAFLSGTFFI   | 60  |
| DB | 1   | MTGTPGAVATRDGEAPERSPPCSYDLTGKVMILGDTGVGKTCFLIOFKDGAFLSGTFFI   | 60  |
| QY | 61  | ATVGTDFRNKVVTVGVRVKLIQIWDTAGQERFVSHTAYYRDAQALLLYDITNKSSFDN    | 120 |
| DB | 61  | ATVGTDFRNKVVTVGVRVKLIQIWDTAGQERFVSHTAYYRDAQALLLYDITNKSSFDN    | 120 |
| QY | 121 | IRANLITEIHEYAQRDVIIMLLGNKADMSSEVIRSEDETGLAREYGVFPFLETSAKTGMNV | 180 |
| DB | 121 | IRANLITEIHEYAQRDVIIMLLGNKADMSSEVIRSEDETGLAREYGVFPFLETSAKTGMNV | 180 |
| QY | 181 | ELAFIAIAKELKYRAGHOADEPFOIRDYVESQKRSCCSFPM                     | 223 |
| DB | 181 | ELAFIAIAKELKYRAGHOADEPFOIRDYVESQKRSCCSFPM                     | 223 |

##### RESULT 2



US-09-794-257-14

Query Match 85.0%; Score 977; DB 10; Length 191;  
Best Local Similarity 100.0%; Pred. No. 5e-91;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| QY | 33  | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 92  |
| Db | 1   | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 60  |
| QY | 93  | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVIIMLLGNKADMSERV  | 152 |
| Db | 61  | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVIIMLLGNKADMSERV  | 120 |
| QY | 153 | IRSEDETAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES  | 212 |
| Db | 121 | IRSEDETAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES  | 180 |
| QY | 213 | QKKRSSCCSFM 223  |     |
| Db | 181 | QKKRSSCCSFM 191  |     |

## RESULT 6

US-10-051-986-3  
; Sequence 3, Application US/10051986  
; Patent No. US20020146770A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; Tang, Y. Tom  
; Lal, Preeti  
; Guegler, Karl J.  
; Corley, Neil C.  
; Patterson, Chandra  
; Batra, Sajeev  
; Baughn, Mariah R.  
; TITLE OF INVENTION: RAS PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051.986  
; FILING DATE: 15-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/766,551  
; FILING DATE: DECEMBER 12, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrione, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 191 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: UCML5701

; CLONE: 1528559  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3 :  
US-10-051-986-3

Query Match 85.0%; Score 977; DB 12; Length 191;  
Best Local Similarity 100.0%; Pred. No. 5e-91;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| QY | 33  | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 92  |
| Db | 1   | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 60  |
| QY | 93  | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVIIMLLGNKADMSERV  | 152 |
| Db | 61  | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVIIMLLGNKADMSERV  | 120 |
| QY | 153 | IRSEDETAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES  | 212 |
| Db | 121 | IRSEDETAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES  | 180 |
| QY | 213 | QKKRSSCCSFM 223  |     |
| Db | 181 | QKKRSSCCSFM 191  |     |

## RESULT 7

US-09-867-550-1812  
; Sequence 1812, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867.550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208.427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1812  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-1812

Query Match 47.6%; Score 547; DB 10; Length 106;  
Best Local Similarity 100.0%; Pred. No. 7.1e-48;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |  |    |
|----|----|--|----|
| QY | 33 | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 92 |
| Db | 1  | MLLGDGVTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDCGVRVKLQIWDTAGQERF | 60 |
| QY | 93 | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVI 138            |    |
| Db | 61 | RSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTHEIHYAQRDVI 106            |    |

## RESULT 8

US-09-764-868-688  
; Sequence 688, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764.868  
; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 688

; LENGTH: 139

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-868-688

Query Match 46.5%; Score 535; DB 9; Length 139;

Best Local Similarity 71.5%; Pred. No. 1.7e-46;

Matches 98; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 84 WDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGN 143

DB 1 WDTAGQERFRSVTHAYYRDAHALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGN 60

QY 144 KADMSSEVIRSEGETLAREYGVPPFLETSKTMNVELAFIAIAKELKYRAGHQADPS 203

DB 61 KVDASHERVVKREDGEKLAKEYGUPFMETSAKTGLNVDLAFIAIAKELKQKRMKAPSEPR 120

QY 204 FOIRDYVESQKRSSCC 220

DB 121 FELHDYVKEGEGASCC 137

#### RESULT 9

US-09-794-257-8

; Sequence 8, Application US/09794257

; Patent No. US20020009804A1

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel

; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1

; FILE REFERENCE: Human G-Proteins

; CURRENT APPLICATION NUMBER: 35800/209285

; PRIOR FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: US/09/794,257

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 207

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-794-257-8

Query Match

Best Local Similarity 43.6%; Score 501; DB 10; Length 207;

Matches 98; Conservative 40; Mismatches 59; Indels 2; Gaps 2;

QY 25 STDLTGKVMMLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIW 84

DB 4 TYDILFKLLIGDSGVGKTCFLRFSEDAF-NTTFISTIGIDFKIRTIELDGGKIKLQIW 62

QY 85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGNK 144

DB 63 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGNK 122

QY 145 ADMSSSERVIRSEGETLAREYGVPPFLETSKTMNVELAFIAIAKELKYRAGHQADPS 203

DB 123 CDMDNRQVSKERGEKLAIDYGIKFLTSKSSANVEEAFFTLARDIMTKLNKRNKNDNS 182

QY 204 FOIRDYVESQKRSSCCSF 222

DB 183 AGAGGVPKRTENRSKTSF 201

#### RESULT 10

US-09-822-860-5

; Sequence 5, Application US/09822860

; Patent No. US20020146795A1

; GENERAL INFORMATION:

; APPLICANT: ZHU, Shiaoqing et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS, NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01214

; CURRENT APPLICATION NUMBER: US/09/822,860

; CURRENT FILING DATE: 2001-04-02

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 190

; TYPE: PRT

; ORGANISM: Discopyge ommata

US-09-822-860-5

Query Match

Best Local Similarity 43.5%; Score 500; DB 10; Length 190;

Matches 96; Conservative 42; Mismatches 51; Indels 2; Gaps 2;

QY 27 DLTGKVMMLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWD 86

DB 1 DLTGKVMMLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWD 59

QY 87 AGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGNKAD 146

DB 60 AGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLGNKAD 119

QY 147 MSSSERVIRSEGETLAREYGVPPFLETSKTMNVELAFIAIAKELKYRAGHQADPSQF- 205

DB 120 MNEKQVSKERGEKLAIDYGIKFLTSKSSINVEEAFFTLARDIMTKLNKKNMNSLQ 179

QY 206 IRDYVESQKR 216

DB 180 AVDLKSPKK 190

#### RESULT 11

US-09-925-300-1571

; Sequence 1571, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Steve Ruben,

; APPLICANT: Steve Ruben,

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 1571

; LENGTH: 218

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-300-1571

Query Match

Best Local Similarity 43.3%; Score 498; DB 10; Length 218;

Matches 99; Conservative 36; Mismatches 63; Indels 6; Gaps 3;

QY 20 PPCSPSYDLTGKVMMLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRV 79

DB 18 PMAKTYDLLFKLLIGDSGVGKTCFLRFSEDAF-NTTFISTIGIDFKIRTVELQGGKI 76

QY 80 KLIQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIM 139

DB 77 KLIQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIM 136

QY 140 LIGNKADMSSEVIRSEGETLAREYGVPPFLETSKTMNVELAFIAIAKELKYRAGHQ 199

Db 137 LLGNKCDMDKRVVPKGGKQIAREHGIRFFETSAKANINIEKAFILAEILRKT--PV 194  
Qy 200 DEPSFQIRDYVES---QKRSSCC 220  
Db 195 KEPNSENVDISSGGVGTGWSKCC 218

## RESULT 12

US-09-834-765-766  
; Sequence 766, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Paris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E. H. Afar  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.60SUI  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 766  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-766

Query Match 43.3%; Score 497.5; DB 10; Length 162;  
Best Local Similarity 55.6%; Pred. No. 1.3e-42;  
Matches 90; Conservative 35; Mismatches 36; Indels 1; Gaps 1;

Qy 31 KVMILGDTGVKTCFLIQKFDGAFSLGTFIATVGIIDFRNKKVTVGVYRVKLIQWDTAGQE 90  
Db 2 KLLIGDSGVKTCVLFKRFSEDAF-NSTFISTIGIDFKIRTIELDKRIKLIQWDTAGQE 60  
Qy 91 RFRSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEVAORDVVMILGNKADMSSE 150  
Db 61 RFTITAYYRGAMGIMLVYDITNEKSFNIRWIRNIEHASADVEMKILGNKCDVNDK 120  
Qy 151 RVIRSEGETLAREYGVPPLETSKAGTMNVELAFILAEILK 192  
Db 121 ROVSKERGERKLDYGIKFKMETSAKANINVENAFFETLARDIK 162

## RESULT 13

US-09-925-302-534  
; Sequence 534, Application US/09925302  
; Patent No. US20020044941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 534  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-302-534

Query Match 42.0%; Score 483; DB 10; Length 246;

Best Local Similarity 44.4%; Pred. No. 6.9e-41;  
Matches 95; Conservative 50; Mismatches 57; Indels 12; Gaps 4;  
Qy 16 PERSPPCS---PSYDLTGKVMILGDTGVKTCFLIQKFDGAFSLGTFIATVGIIDFRNKKV 72  
Db 35 PPPAPPVSAMAKAYDHLFKLLIGDSGVKTCILIRAEADNF-NTYISTIGIDFKRTV 93  
Qy 73 TVDGVYRVKLIQWDTAGQERFSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEYA 132  
Db 94 DIEGKIKLQWDTAGQERFKTITAYYRGAMGILVYDITDEKSFENIQNMKSIRENA 153  
Qy 133 QRDVVMILGNKADMSSEVIRSEGETLAREYGVPPLETSKAGTMNVELAFILAEILK 192  
Db 154 SAGVERLLGNKCDMEARKVKQEQADKLAREHGIRFFETSAKANINIEKAFILAEIL 213  
Qy 193 Y-----RAGHOADPSFQIRDYVESQKRSSCCS 221  
Db 214 LKSGRSGNGNKPPS---TDLKTCDDKKNKCS 244

## RESULT 14

US-09-967-736-8  
; Sequence 8, Application US/09967736  
; Patent No. US20020103340A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; Lal, Preeti  
; Corley, Neil C.  
; Shah, Purvi  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER: READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/967,736  
; FILING DATE: 28-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/154,602  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0367 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 57006  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-967-736-8

Query Match 41.0%; Score 472; DB 10; Length 201;  
Best Local Similarity 43.1%; Pred. No. 6.8e-40;  
Matches 87; Conservative 51; Mismatches 58; Indels 6; Gaps 3;

Qy 23 SPSTDLTGKVMILGDTGVKTCFLIQKFDGAFSLGTFIATVGIIDFRNKKVTVGVYRVKLIQ 82

Db 2 NPEYDYLKLLIGDSGVGKSCLLRFADDTY-TESYISTIGVDFKRTIELDGTIKLQ 60  
QY 83 IWDTAQGRFRSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTHIHEYAQRDVVIMLG 142  
Db 61 IWDTAQGRFRSVTHAYYRGAHGIIVYDVTDQESYANVKQLQELIDRYASENVNKLVG 120  
QY 143 NKADMSRVRIRSEGETLAREYGVFPFLETSAGTGMNVELAFALAKELKYRAGHQA--- 199  
Db 121 NKSDLTTRKKVNDTTAKEFADSLGVFPFLETSAGTGMNVELAFALAKELKYRAGHQA--- 180  
QY 200 -DEPSFOIRDYVESQKRSSCC 220  
Db 181 GERPNLKI-DSTPVKASGGCC 201

RESULT 15

US-10-102-806-466  
; Sequence 466, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103P1C1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 466  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-102-806-466

Query Match 40.8%; Score 469.5; DB 9; Length 224;  
Best Local Similarity 40.2%; Pred. No. 1.4e-39;  
Matches 90; Conservative 56; Mismatches 65; Indels 13; Gaps 5;  
QY 8 VATRGE-----APERSPPC--SPSYDLTGKVMILGDTGVGKTCFLIOFKDGAFLSGTFI 60  
Db 3 ILEREAEOSRLGATERAAAAANPEYDYLKLLIGDSGVGKSCLLRFADDTY-TESYI 61  
QY 61 ATVGIDFRNKVTVVGVKVLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFDN 120  
Db 62 STIGVDFKRTIELDGTIKLQIWDTAGQERFRSVTHAYYRGAHGIIVYDVTDQESYAN 121  
QY 121 IRAWLTHIHEYAQRDVVIMLGKADMSRVRIRSEGETLAREYGVFPFLETSAGTGMNV 180  
Db 122 VQWLQELIDRYASENVNKLVGKNSDLTTKKVVDNTTAKFADSLGIPFLETSAGTGMNV 181  
QY 181 ELAFLALAKELKYRAGHQA---DEPSFOIRDYVESQKRSSCC 220  
Db 182 EQAFMTMAEIKKRMGPGAASGERPNLKI-DSTPVKAGGGCC 224

Search completed: June 18, 2003, 15:56:57  
Job time : 49 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:48:03 ; Search time 28 seconds  
(without alignments)  
234.333 Million cell updates/sec

Title: US-09-817-199B-2

Perfect score: 1150

Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 977   | 85.0        | 191    | 4     | US-09-075-454-3   |
| 2          | 745   | 64.8        | 190    | 3     | US-08-824-873-3   |
| 3          | 745   | 64.8        | 190    | 3     | US-09-198-184-3   |
| 4          | 666   | 57.9        | 190    | 2     | US-08-824-873-1   |
| 5          | 666   | 57.9        | 190    | 3     | US-09-198-184-1   |
| 6          | 513.5 | 44.7        | 207    | 2     | US-08-824-873-4   |
| 7          | 513.5 | 44.7        | 207    | 3     | US-09-198-184-4   |
| 8          | 504   | 43.8        | 205    | 2     | US-08-531-525-25  |
| 9          | 504   | 43.8        | 205    | 2     | US-08-718-270A-25 |
| 10         | 486   | 42.3        | 207    | 2     | US-08-531-525-35  |
| 11         | 486   | 42.3        | 207    | 2     | US-08-718-270A-35 |
| 12         | 478.5 | 41.6        | 198    | 2     | US-08-531-525-51  |
| 13         | 478.5 | 41.6        | 198    | 2     | US-08-718-270A-51 |
| 14         | 478.5 | 41.6        | 215    | 2     | US-08-531-525-10  |
| 15         | 478.5 | 41.6        | 215    | 2     | US-08-718-270A-10 |
| 16         | 472   | 41.0        | 201    | 2     | US-08-916-901-8   |
| 17         | 472   | 41.0        | 201    | 4     | US-09-154-602-8   |
| 18         | 471   | 41.0        | 202    | 2     | US-08-531-525-14  |
| 19         | 471   | 41.0        | 202    | 2     | US-08-718-270A-14 |
| 20         | 468   | 40.7        | 201    | 2     | US-08-916-901-3   |
| 21         | 468   | 40.7        | 201    | 4     | US-09-154-602-3   |
| 22         | 458   | 39.8        | 201    | 2     | US-08-531-525-13  |
| 23         | 458   | 39.8        | 201    | 2     | US-08-718-270A-13 |
| 24         | 450.5 | 39.2        | 194    | 2     | US-08-531-525-34  |
| 25         | 450.5 | 39.2        | 194    | 2     | US-08-718-270A-34 |
| 26         | 448.5 | 39.0        | 212    | 4     | US-09-399-913-67  |
| 27         | 446.5 | 38.8        | 212    | 2     | US-08-531-525-18  |

Sequence 18, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 12, Appl  
Sequence 12, Appl  
Sequence 8, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 52, Appl  
Sequence 52, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 11, Appl  
Sequence 11, Appl

28 446.5 38.8 212 2 US-08-718-270A-18  
29 426.5 37.1 210 2 US-08-531-525-16  
30 426.5 37.1 210 2 US-08-718-270A-16  
31 410 35.7 208 2 US-08-531-525-17  
32 410 35.7 208 2 US-08-718-270A-17  
33 409.5 35.6 218 2 US-08-531-525-19  
34 409.5 35.6 218 2 US-08-718-270A-19  
35 401 34.9 88 3 US-08-665-259-12  
36 391.5 34.0 203 2 US-08-766-551-8  
37 391.5 34.0 203 2 US-08-531-525-15  
38 389.5 33.9 208 2 US-08-718-270A-15  
39 389.5 33.9 208 2 US-08-531-525-15  
40 381 33.1 214 2 US-08-531-525-52  
41 381 33.1 214 2 US-08-718-270A-52  
42 364 31.7 203 2 US-08-531-525-21  
43 364 31.7 203 2 US-08-718-270A-21  
44 361 31.4 213 2 US-08-531-525-11  
45 361 31.4 213 2 US-08-718-270A-11

#### ALIGNMENTS

RESULT 1  
US-09-075-454-3  
; Sequence 3, Application US/09075454  
; Patent No. 6391580  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Lal, Preeti  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Batra, Sajeev  
; APPLICANT: Haughn, Mariah R.  
; TITLE OF INVENTION: RAS PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,454  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/766,551  
; FILING DATE: DECEMBER 12, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Certone, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/POCKET NUMBER: PF-0168-1 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 191 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: UCMCL5T01

```
; CLONE: 1528559
US-09-075-454-3
Query Match      85.0%; Score 977; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 3.9e-101;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQFDKGAFLSGTFTIATVGIDFRNKVVTVDGVRVKLIQIWDTAGQERF 92
Db 1 MLVGDSGVGKTCCLLVRFKDGAFAGTFTSTVGIDFRNKVLDVDMKVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152
Db 61 RSVTHAYYRDAHALLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120
QY 153 IRSEGETLAREYGVPFLETSKATGMNVELAFALAKELKYRAGHQADEPSPQIRDYVES 212
Db 121 IRSEGETLAREYGVPFLETSKATGMNVELAFALAKELKYRAGHQADEPSPQIRDYVES 180
QY 213 QKRRSCCSFM 223
Db 181 QKRRSCCSFM 191

RESULT 2
US-08-824-873-3
; Sequence 3, Application US/08824873
; Patent No. 5843717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; APPLICATION DATA:
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 190 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 619734
US-08-824-873-3
Query Match      64.8%; Score 745; DB 2; Length 190;
Best Local Similarity 73.4%; Pred. No. 3e-75;
Matches 138; Conservative 27; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQFDKGAFLSGTFTIATVGIDFRNKVVTVDGVRVKLIQIWDTAGQERF 92
Db 1 MLVGDSGVGKTCCLLVRFKDGAFAGTFTSTVGIDFRNKVLDVDMKVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152
Db 61 RSVTHAYYRDAHALLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120
QY 153 IRSEGETLAREYGVPFLETSKATGMNVELAFALAKELKYRAGHQADEPSPQIRDYVES 212
Db 121 IRSEGETLAREYGVPFLETSKATGMNVELAFALAKELKYRAGHQADEPSPQIRDYVES 180
QY 213 QKRRSCCSFM 223
Db 181 QKRRSCCSFM 191

RESULT 3
US-09-198-184-3
; Sequence 3, Application US/09198184
; Patent No. 6010859
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/824,873
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 190 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 619734
US-09-198-184-3
Query Match      64.8%; Score 745; DB 3; Length 190;
Best Local Similarity 73.4%; Pred. No. 3e-75;
Matches 138; Conservative 27; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQFDKGAFLSGTFTIATVGIDFRNKVVTVDGVRVKLIQIWDTAGQERF 92
Db 1 MLVGDSGVGKTCCLLVRFKDGAFAGTFTSTVGIDFRNKVLDVDMKVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152
```

Db 61 RSVTHAYRRDAHALLLDYDTNKDSFNDNIQAWLTFIEYQAQDDVVLMLGNKKVDSQERV 120  
QY 153 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFOIRDYVES 212  
Db 121 VKREDGEKLAKEYGLPMTSAKSGNLVDAFLTAIAKELKQSTKAPSEPRFLHDYVVR 180  
QY 213 QKRSSCC 220  
Db 181 EGRGASCC 188

## RESULT 4

US-08-824-873-1  
; Sequence 1, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; APPLICATION DATA:  
; FILING DATE: Filled Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 190 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT04  
; CLONE: 738957  
US-08-824-873-1

Query Match 57.9%; Score 666; DB 2; Length 190;  
Best Local Similarity 67.08; Pred. No. 2e-66;  
Matches 126; Conservative 25; Mismatches 37; Indels 0; Gaps 0;  
QY 33 MLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKNVTVGVRVKLQIWDTAGQERF 92  
Db 1 MLVGDGSGVGTCLLGAIQGWCFCPGDHLHRSIDFRNKNVLDVGVKVKLQMDWDTAGQERF 60  
QY 93 RSVTHAYRRDAHALLLDYDTNKDSFNDNIQAWLTFIEYQAQDDVVLMLGNKKVDSQERV 152  
Db 61 RSVTHAYRRDAHALLLDYDTNKDSFNDNIQAWLTFIEYQAQDDVVLMLGNKKVDSQERV 120  
QY 153 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFOIRDYVES 212  
Db 121 VKREDGEKLAKEYGLPMTSAKSGNLVDAFLTAIAKELKQSTKAPSEPRFLHDYVVR 180

QY 213 QKRSSCC 220  
Db 181 EGRGASCC 188

## RESULT 5

US-09-198-184-1  
; Sequence 1, Application US/09198184  
; Patent No. 6010859  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09198,184  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,873  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 190 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT04  
; CLONE: 738957  
US-09-198-184-1

Query Match 57.9%; Score 666; DB 3; Length 190;  
Best Local Similarity 67.08; Pred. No. 2e-66;  
Matches 126; Conservative 25; Mismatches 37; Indels 0; Gaps 0;  
QY 33 MLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKNVTVGVRVKLQIWDTAGQERF 92  
Db 1 MLVGDGSGVGTCLLGAIQGWCFCPGDHLHRSIDFRNKNVLDVGVKVKLQMDWDTAGQERF 60  
QY 93 RSVTHAYRRDAHALLLDYDTNKDSFNDNIQAWLTFIEYQAQDDVVLMLGNKKVDSQERV 152  
Db 61 RSVTHAYRRDAHALLLDYDTNKDSFNDNIQAWLTFIEYQAQDDVVLMLGNKKVDSQERV 120  
QY 153 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFOIRDYVES 212  
Db 121 VKREDGEKLAKEYGLPMTSAKSGNLVDAFLTAIAKELKQSTKAPSEPRFLHDYVVR 180  
QY 213 QKRSSCC 220  
Db 181 EGRGASCC 188

## RESULT 6

US-08-824-873-4  
 ; Sequence 4, Application US/08824873  
 ; Patent No. 5843717  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/824,873  
 ; FILING DATE: Filed Herewith  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0240 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 207 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 234746  
 ; US-08-824-873-4

Query Match 44.7%; Score 513.5; DB 2; Length 207;  
 Best Local Similarity 50.0%; Pred. No. 2.4e-49;  
 Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;  
 QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFIATVGVDFRNKVVYDGVVRVQLQIW 84  
 Db :|| :||:||||| :|| :||:||||| :||:||||| :||:||||| :||:|||||  
 4 TYDYLKLLIGDSGVGKTCVLFREFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLIQIW 62  
 QY 85 DTAGQERFSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTIETHEYAQRDVVIMLGNK 144  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 63 DTAGQERFTITAYRGAMGIMLVYDITNEKSFNIRNIRNIEEHASADVERKMILGNK 122  
 QY 145 ADMSSERVIRSEGETLAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQAD---- 200  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 123 CDVNDKRVQSKERGEKALDYIKFPMETSAKANINVENAFTLARDIRAKMDKLEGNP 182  
 QY 201 EPSFQIRDYVESQKRSS 218  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 183 QGSNOGVKITPPQQRSS 200

## RESULT 7

US-09-198-184-4  
 ; Sequence 4, Application US/09198184  
 ; Patent No. 6010859  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Guegler, Karl  
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/198,184  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/824,873  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0240 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 207 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 234746  
 ; US-09-198-184-4

Query Match 44.7%; Score 513.5; DB 3; Length 207;  
 Best Local Similarity 50.0%; Pred. No. 2.4e-49;  
 Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;  
 QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFIATVGVDFRNKVVYDGVVRVQLQIW 84  
 Db :|| :||:||||| :|| :||:||||| :||:||||| :||:||||| :||:|||||  
 4 TYDYLKLLIGDSGVGKTCVLFREFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLIQIW 62  
 QY 85 DTAGQERFSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLTIETHEYAQRDVVIMLGNK 144  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 63 DTAGQERFTITAYRGAMGIMLVYDITNEKSFNIRNIRNIEEHASADVERKMILGNK 122  
 QY 145 ADMSSERVIRSEGETLAREYGVFPFLETSKATGMNVELAFIAIAKELKYRAGHQAD---- 200  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 123 CDVNDKRVQSKERGEKALDYIKFPMETSAKANINVENAFTLARDIRAKMDKLEGNP 182  
 QY 201 EPSFQIRDYVESQKRSS 218  
 Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||  
 183 QGSNOGVKITPPQQRSS 200

## RESULT 8

US-08-531-525-25  
 ; Sequence 25, Application US/08531525  
 ; Patent No. 5840683  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hlavka, Joseph J.  
 ; APPLICANT: Pincus, Matthew R.  
 ; APPLICANT: No. 5840683le, John F.  
 ; APPLICANT: Abajian, Henry B.  
 ; APPLICANT: Kende, Andrew S.  
 ; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
 ; TITLE OF INVENTION: of P21 Ras



ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-531-525-35

Query Match 42.3%; Score 486; DB 2; Length 207;  
Best Local Similarity 48.1%; Pred. No. 2.8e-46;  
Matches 101; Conservative 44; Mismatches 49; Indels 16; Gaps 6;  
QY 25 SYDLTGKVMVLGDTGVGKTCFLIQKDGAFUSGTFTIATVGDIFRNKVVTVDSVRVKLIQIW 84  
DB 3 TYDYLKLLIGDSGVGKTCFLFRSEDAF-NTTFTISTIGIDFKIRTVELDGKKIKLIQIW 61  
QY 85 DTAGQRRFRSTHAYYRDAQAALLLYDITNKSSFDNIRAWLTIETHEYAQRDVVIMLGNK 144  
DB 62 DTAGQRRFTIT-AYIRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
QY 145 ADMSSERVIRSEDEGTALREYGVPELTSAGTGMNVELAFIAIAKELKYRAGHQADEPSF 204  
DB 120 CDMEKROVSKERGEKLAIDYGIKLETSKASSINVEEAFITLARDINTKLKKNENSL 179  
QY 205 QIRDYVE-----SOKKR--SSCCSFM 223  
DB 180 Q--EAVDKLKSPKPKPSOKKQLSPRCSLL 207

RESULT 11  
US-08-718-270A-35  
Sequence 35, Application US/08/18270A  
Patent No. 5910478  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: NO. 5910478le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado

COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-718-270A-35

Query Match 42.3%; Score 486; DB 2; Length 207;  
Best Local Similarity 48.1%; Pred. No. 2.8e-46;  
Matches 101; Conservative 44; Mismatches 49; Indels 16; Gaps 6;  
QY 25 SYDLTGKVMVLGDTGVGKTCFLIQKDGAFUSGTFTIATVGDIFRNKVVTVDSVRVKLIQIW 84  
DB 3 TYDYLKLLIGDSGVGKTCFLFRSEDAF-NTTFTISTIGIDFKIRTVELDGKKIKLIQIW 61  
QY 85 DTAGQRRFRSTHAYYRDAQAALLLYDITNKSSFDNIRAWLTIETHEYAQRDVVIMLGNK 144  
DB 62 DTAGQRRFTIT-AYIRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
QY 145 ADMSSERVIRSEDEGTALREYGVPELTSAGTGMNVELAFIAIAKELKYRAGHQADEPSF 204  
DB 120 CDMEKROVSKERGEKLAIDYGIKLETSKASSINVEEAFITLARDINTKLKKNENSL 179  
QY 205 QIRDYVE-----SOKKR--SSCCSFM 223  
DB 180 Q--EAVDKLKSPKPKPSOKKQLSPRCSLL 207

RESULT 12  
US-08-531-525-51  
Sequence 51, Application US/08/531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: NO. 5840683le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201

```

; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
;
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
;
; US-08-531-525-51

```

Query Match 41.6%; Score 478.5; DB 2; Length 198;

Best Local Similarity 48.7%; Pred. No. 1.8e-45;

Matches 97; Conservative 36; Mismatches 59; Indels 7; Gaps 4;

```

Qy 25 SYDLTGKVMILGDTGVGKTCFLQKDGAFSLGTFIATVGDIFRNKVVYDGVVRKLIQW 84
Db 4 TYDLLEKLLIGDSGVGKTCVLFKFSDDAF-NTTFI-SIGIDFKIKTVLQGGKIKLIQW 61
Qy 85 DTAGQERFSRTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHEYQAQRDVVIMLGK 144
Db 62 DTAGQERFHTTTSYRGANGIMLVYDITNKSFENISKWLNRIDEHANEDVERMLGNK 121
Qy 145 ADMSSERVISEDEGTAREYGVPLFETSAKTMNVELAFIAIAKELKYRAGHQADPSF 204
Db 122 CDMDKRVVPKGGKGEQIAREHGIRFFETSAKVNIINIEKAFITLAEDILRKT--PVKEPNS 179
Qy 205 QIRDYVES---QKRRSICC 220
Db 180 ENVDISSGGVGTGWKSKCC 198

```

#### RESULT 13

US-08-718-270A-51

; Sequence 51, Application US/08718270A

; Patent No. 5910478

; GENERAL INFORMATION:

; APPLICANT: Hlavka, Joseph J.

; APPLICANT: Pincus, Matthew R.

; APPLICANT: No. 5910478le, John F.

; APPLICANT: Abajian, Henry B.

; APPLICANT: Kende, Andrew S.

; TITLE OF INVENTION: Peptidomimetics Inhibiting

; TITLE OF INVENTION: the Oncogenic Action of P21 Ras

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

; STREET: 5370 Manhattan Circle, Suite 201

; CITY: Boulder

; STATE: Colorado

; COUNTRY: US

; ZIP: 80303

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COUNTRY: US
; ZIP: 80303
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
;
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
;
; US-08-718-270A-51

```

Query Match 41.6%; Score 478.5; DB 2; Length 198;

Best Local Similarity 48.7%; Pred. No. 1.8e-45;

Matches 97; Conservative 36; Mismatches 59; Indels 7; Gaps 4;

```

Qy 25 SYDLTGKVMILGDTGVGKTCFLQKDGAFSLGTFIATVGDIFRNKVVYDGVVRKLIQW 84
Db 4 TYDLLEKLLIGDSGVGKTCVLFKFSDDAF-NTTFI-SIGIDFKIKTVLQGGKIKLIQW 61
Qy 85 DTAGQERFSRTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHEYQAQRDVVIMLGK 144
Db 62 DTAGQERFHTTTSYRGANGIMLVYDITNKSFENISKWLNRIDEHANEDVERMLGNK 121
Qy 145 ADMSSERVISEDEGTAREYGVPLFETSAKTMNVELAFIAIAKELKYRAGHQADPSF 204
Db 122 CDMDKRVVPKGGKGEQIAREHGIRFFETSAKVNIINIEKAFITLAEDILRKT--PVKEPNS 179
Qy 205 QIRDYVES---QKRRSICC 220
Db 180 ENVDISSGGVGTGWKSKCC 198

```

#### RESULT 14

US-08-531-525-10

; Sequence 10, Application US/08531525

; Patent No. 5840683

; GENERAL INFORMATION:

; APPLICANT: Hlavka, Joseph J.

; APPLICANT: Pincus, Matthew R.

; APPLICANT: No. 5840683le, John F.

; APPLICANT: Abajian, Henry B.

; APPLICANT: Kende, Andrew S.

; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action

; TITLE OF INVENTION: of P21 Ras

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenlee and Winner, P.C.

; STREET: 5370 Manhattan Circle, Suite 201

; CITY: Boulder

; STATE: Colorado

```

; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; US-08-531-525-10

Query Match 41.6%; Score 478.5; DB 2; Length 215;
Best Local Similarity 45.1%; Pred. No. 2e-45;
Matches 96; Conservative 50; Mismatches 54; Indels 13; Gaps 6;

QY 19 SPP--CSPSYDITGKVMILGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDG 76
DB 3 APPARADYDYLLKLLIGDSGVKSCLLLRFSGSGFTT-SFITTIGIDFKIRTEILDG 61
QY 77 VVKLIQIWDTAGQERFSTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
DB 62 KRIKLIQIWDTAGQER-RITTTAYRGAMGILLVYDVTDESSFNIRNIRNIEQHASDNV 120
QY 137 VMILGNKADM--SSERVISEDGETLAREYGVFPFLETSAKTMGNVELAFALAKELKYR- 194
DB 121 NKILVGNKADMDESKRAVPTAKQALADEYGIKFETSAKTNLNLNVEEYFFSIGRIDIKQRL 180
QY 195 --AGHQADEPSFOIRDYVES-----QKKRSSCC 220
DB 181 SDTDSRAEPATIKISQTDQAAGAGQATQKSACC 213

RESULT 15
US-08-718-270A-10
; Sequence 10, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718/270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; US-08-718-270A-10

Query Match 41.6%; Score 478.5; DB 2; Length 215;
Best Local Similarity 45.1%; Pred. No. 2e-45;
Matches 96; Conservative 50; Mismatches 54; Indels 13; Gaps 6;

QY 19 SPP--CSPSYDITGKVMILGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVDG 76
DB 3 APPARADYDYLLKLLIGDSGVKSCLLLRFSGSGFTT-SFITTIGIDFKIRTEILDG 61
QY 77 VVKLIQIWDTAGQERFSTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
DB 62 KRIKLIQIWDTAGQER-RITTTAYRGAMGILLVYDVTDESSFNIRNIRNIEQHASDNV 120
QY 137 VMILGNKADM--SSERVISEDGETLAREYGVFPFLETSAKTMGNVELAFALAKELKYR- 194
DB 121 NKILVGNKADMDESKRAVPTAKQALADEYGIKFETSAKTNLNLNVEEYFFSIGRIDIKQRL 180
QY 195 --AGHQADEPSFOIRDYVES-----QKKRSSCC 220
DB 181 SDTDSRAEPATIKISQTDQAAGAGQATQKSACC 213

Search completed: June 18, 2003, 15:56:01
Job time : 29 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:36:32 ; Search time 87 Seconds  
(without alignments)  
528.144 Million cell updates/sec

Title: US-09-817-199B-2

Perfect score: 1150

Sequence: 1 MTCTGAVATRDGEAPERSP.....FQIRDYVESOKKRSKSCFSM 223

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvrius:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 697   | 60.6        | 666    | 5 Q9VP48  | Q9vp48 drosophila   |
| 2          | 648.5 | 56.4        | 388    | 5 Q9SR32  | Q9sr32 drosophila   |
| 3          | 533   | 46.3        | 196    | 5 O02046  | O02046 caenorhabdi  |
| 4          | 521.5 | 45.3        | 207    | 11 Q8VCF6 | Q8vcf6 mus musculus |
| 5          | 514.5 | 44.7        | 202    | 10 Q9SXT5 | Q9sxt5 cicer ariet  |
| 6          | 513.5 | 44.7        | 202    | 10 Q08153 | Q08153 pisum sativ  |
| 7          | 510   | 44.3        | 216    | 10 Q49844 | Q49844 daucus caro  |
| 8          | 509.5 | 44.3        | 214    | 10 Q40218 | Q40218 lotus japon  |
| 9          | 509   | 44.3        | 203    | 10 Q949E2 | Q949e2 oryza sativ  |
| 10         | 509   | 44.3        | 216    | 10 Q40215 | Q40215 lotus japon  |
| 11         | 509   | 44.3        | 218    | 10 Q9SE91 | Q9se91 arabidopsis  |
| 12         | 508.5 | 44.2        | 204    | 5 O15971  | O15971 drosophila   |
| 13         | 507.5 | 44.1        | 202    | 10 Q40203 | Q40203 lotus japon  |
| 14         | 507   | 44.1        | 216    | 10 Q9LZD4 | Q9lzd4 arabidopsis  |
| 15         | 507   | 44.1        | 216    | 10 Q8VWF9 | Q8vwf9 nicotiana t  |
| 16         | 506   | 44.0        | 203    | 10 Q9M7P5 | Q9m7p5 capsicum an  |

|    |       |      |     |    |        |                    |
|----|-------|------|-----|----|--------|--------------------|
| 17 | 506   | 44.0 | 216 | 10 | Q40177 | Q40177 lycopersico |
| 18 | 505.5 | 44.0 | 201 | 5  | Q94148 | Q94148 caenorhabdi |
| 19 | 505   | 43.9 | 203 | 10 | Q24112 | Q24112 nicotiana p |
| 20 | 505   | 43.9 | 216 | 10 | Q8W3J4 | Q8w3j4 nicotiana t |
| 21 | 504   | 43.8 | 216 | 10 | Q8W3J3 | Q8w3j3 nicotiana t |
| 22 | 503   | 43.7 | 203 | 10 | Q41340 | Q41340 lycopersico |
| 23 | 502   | 43.7 | 216 | 10 | Q9FJF1 | Q9fjf1 arabidopsis |
| 24 | 501.5 | 43.6 | 203 | 10 | Q8RU63 | Q8ru63 oryza sativ |
| 25 | 501   | 43.6 | 203 | 10 | Q94027 | Q94027 arabidopsis |
| 26 | 500   | 43.5 | 216 | 10 | Q24466 | Q24466 arabidopsis |
| 27 | 499.5 | 43.4 | 216 | 10 | Q8W3J2 | Q8w3j2 nicotiana t |
| 28 | 499   | 43.4 | 212 | 10 | Q40217 | Q40217 lotus japon |
| 29 | 498   | 43.3 | 203 | 10 | Q41338 | Q41338 lycopersico |
| 30 | 498   | 43.3 | 216 | 10 | Q9SWV8 | Q9swv8 lycopersico |
| 31 | 497.5 | 43.3 | 202 | 10 | Q9FPJ4 | Q9fpj4 arabidopsis |
| 32 | 497.5 | 43.3 | 202 | 10 | Q08155 | Q08155 pisum sativ |
| 33 | 497.5 | 43.3 | 215 | 10 | Q41023 | Q41023 pisum sativ |
| 34 | 496.5 | 43.2 | 202 | 10 | Q9SEH3 | Q9seh3 arabidopsis |
| 35 | 495   | 43.0 | 215 | 10 | Q40219 | Q40219 lotus japon |
| 36 | 495   | 43.0 | 216 | 10 | Q41024 | Q41024 pisum sativ |
| 37 | 490.5 | 42.7 | 215 | 10 | Q41022 | Q41022 pisum sativ |
| 38 | 489   | 42.5 | 222 | 10 | Q96362 | Q96362 brassica ca |
| 39 | 488   | 42.4 | 203 | 10 | Q08154 | Q08154 pisum sativ |
| 40 | 488   | 42.4 | 203 | 10 | Q40569 | Q40569 nicotiana t |
| 41 | 487.5 | 42.4 | 202 | 10 | Q8W458 | Q8w458 arabidopsis |
| 42 | 487.5 | 42.4 | 207 | 5  | O18338 | O18338 drosophila  |
| 43 | 487   | 42.3 | 200 | 4  | Q9H0T3 | Q9h0t3 homo sapien |
| 44 | 486   | 42.3 | 202 | 10 | Q9ZRH6 | Q9zrh6 petunia hyb |
| 45 | 484.5 | 42.1 | 215 | 10 | Q41061 | Q41061 pisum sativ |

#### ALIGNMENTS

#### RESULT 1

|        |   |   |              |      |         |
|--------|---|---|--------------|------|---------|
| Q9VP48 | ID  | Q9VP48                                  | PRELIMINARY; | PRT; | 666 AA. |
| AC     | Q9VP48  |   |              |      |         |
| DT     | 01-MAY-2000   | (TREMBLrel. 13, Created)                |              |      |         |
| DT     | 01-MAY-2000   | (TREMBLrel. 13, Last sequence update)   |              |      |         |
| DT     | 01-JUN-2002   | (TREMBLrel. 21, Last annotation update) |              |      |         |
| DE     | CG7605  | protein.                                |              |      |         |
| GN     | CG7605  |   |              |      |         |
| OS     | Drosophila melanogaster   | (Fruit fly).                            |              |      |         |
| OC     | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;           |   |              |      |         |
| OC     | Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;   |   |              |      |         |
| OC     | Ephydroidea; Drosophilidae; Drosophila.                                 |   |              |      |         |
| OX     | NCBI_TaxID=7227;  |   |              |      |         |
| RN     | [1]   |   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |   |              |      |         |
| RC     | STRAIN=BERKELEY;  |   |              |      |         |
| RX     | MEDLINE=20196006; PubMed=10731132;                                      |   |              |      |         |
| RA     | Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,         |   |              |      |         |
| RA     | Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,       |   |              |      |         |
| RA     | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,     |   |              |      |         |
| RA     | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,           |   |              |      |         |
| RA     | Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M.R., Pfeiffer B.D.,  |   |              |      |         |
| RA     | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,   |   |              |      |         |
| RA     | Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,   |   |              |      |         |
| RA     | Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,     |   |              |      |         |
| RA     | Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,        |   |              |      |         |
| RA     | Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,           |   |              |      |         |
| RA     | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,   |   |              |      |         |
| RA     | Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,           |   |              |      |         |
| RA     | de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,       |   |              |      |         |
| RA     | Dodson K.J., Dup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., |   |              |      |         |
| RA     | Dubin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,    |   |              |      |         |
| RA     | Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,       |   |              |      |         |
| RA     | Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,            |   |              |      |         |
| RA     | Harris N.I., Harvey D., Helman T.J., Hernandez J.R., Houck J.,          |   |              |      |         |
| RA     | Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,           |   |              |      |         |
| RA     | Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  |   |              |      |         |
| RA     | Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,        |   |              |      |         |

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinekt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; AE003594; AAF51708.1; -;  
 DR HSP; P05713; 3RAB.  
 DR FlyBase; FBgn0037072; CG7605.  
 DR InterPro; IPR003579; GTPase\_Rab.  
 DR InterPro; IPR001806; Ras\_transfmrng.  
 DR InterPro; IPR002078; Sig54\_interact.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 2.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 666 AA; 72746 MW; 153558368B316AD5 CRC64;

Query Match 60.6%; Score 697; DB 5; Length 666;  
 Best Local Similarity 59.8%; Pred. No. 1.5e-55;  
 Matches 137; Conservative 33; Mismatches 49; Indels 10; Gaps 3;  
 3 GTPGAVAT---RDGEA-----PERSPPCSPSYDITGKVLMLGDTGVCCTFLIOFKDGA 53  
 433 GSEGSATLCKNAGRALINMTSSKAPEEEDFDGKVMIGDSGVGKTSLLIRFRDGR 492  
 54 FLGSGFIATVGDIFRNKVVVDGVRVKLIQWDTAGQFRSVTHAYYRDAQALLLYDIT 113  
 493 YVPSYFLSVGIDFRNKVVVDGVRVKLIQWDTAGQFRSVTHAYYRDAHALLLYDVT 552  
 114 NKSSFDNIRAWLTHEIYAQRDVMILGNKADMS-SERVIRSEGETLAREYGVPPLET 172  
 553 NKTYYDNIRAWLGEIREYAEQDVIVLIGNKADCSGSEKQVKREDGERLGRHNPFFMET 612  
 173 SAKTGMNVELAFALAKELKYRAGHQADPSFOIRDYVESQKKRSSCCS 221  
 613 SAKTGLNVELSTAVARQLKSRGYEHGDDGKENVHDFVRDNTKARSVCA 661

RESULT 2  
 Q95R32 PRELIMINARY; PRT; 388 AA.  
 AC Q95R32  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE GH21984p.  
 GN CG7605.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,

RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY061826; AAL27637.1; -;  
 DR FlyBase; FBgn0037072; CG7605.  
 DR InterPro; IPR001806; Ras\_transfmrng.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 KW GTP-binding.  
 SQ SEQUENCE 388 AA; 43065 MW; D57C5244CA53492F CRC64;

Query Match 56.4%; Score 648.5; DB 5; Length 388;  
 Best Local Similarity 65.1%; Pred. No. 2e-51;  
 Matches 125; Conservative 25; Mismatches 41; Indels 1; Gaps 1;  
 31 KVMLLGDTGVGKTCFLIOFKDGAFLSGTFTATVGTIDFRNKVTVTVGVVRVKLIQWDTAGQE 90  
 192 KTIILGDSGVGKTSFLVKNYVTEGFRGSGFATVGTALTNKVVVDGTRVKLIQWDTAGQE 251  
 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIYAQRDVMILGNKADMS-S 149  
 252 RFRSVTHAYYRDAHALLLYDVTNKTYYDNIRAWLGEIREYAEQDVIVLIGNKADCSGS 311  
 150 ERVIRSEGETLAREYGVPPLETSAKTMNVELAFALAKELKYRAGHQADPSFOIRDY 209  
 312 ERQVKREDGERLGRHNPFFMETSAKTMNVELSTAVARQLKSRGYEHGDDGGRFNVHDF 371  
 210 VESQKKRSSCCS 221  
 372 VRDNTKARSVCA 383

RESULT 3  
 Q02046 PRELIMINARY; PRT; 196 AA.  
 AC Q02046  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE W01H2.3 protein.  
 GN W01H2.3.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkison-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Minx P., Wohldmann P.;  
 RT "The sequence of *C. elegans* cosmid W01H2.";  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;

|                       |   |  |                                      |
|-----------------------|---|--|--------------------------------------|
| QY                    | 85  | DTAQGRFRSRVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFTEIHEYAQRDVVIMILGNK | 144                                  |
| Db                    | 63  | DTAQGRFRFTITAYYRGAMGIMLVYDITNEKSFNDINRWIRNIEEHASADVEKMILGNK    | 122                                  |
| QY                    | 145   | ADMSSRSVRSEDETLAREYGVFFLETSAKTGMNVELAFATAKEIKYR-----AGHQA      | 199                                  |
| Db                    | 123   | CVYNDKRVQSKERGEKLADYGIKFMETSAKANINVENAFPTLARDIKAKMDKKLGGNSP    | 182                                  |
| QY                    | 200   | DEPSFOIRDYVESQKKRSS-CSIEM                                      | 223                                  |
| Db                    | 183   | QGSHGKVIYEQKRTSFFRCSL  | 207                                  |
| RESULT 5              |   |  |                                      |
| Q9SXT5                |   | PRELIMINARY;   | PRT; 202 AA.                         |
| ID                    | Q9SXT5  |  |                                      |
| AC                    | Q9SXT5  |  |                                      |
| DT                    | 01-MAY-2000 (TEMBLrel. 13, Created)                                   |  |                                      |
| DT                    | 01-MAY-2000 (TEMBLrel. 13, Last sequence update)                      |  |                                      |
| DT                    | 01-JUN-2002 (TEMBLrel. 21, Last annotation update)                    |  |                                      |
| DE                    | Rab-type small GTP-binding protein.                                   |  |                                      |
| DE                    | Cicer arietinum (Chickpea) (Carbanzo).                                |  |                                      |
| OC                    | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;    |  |                                      |
| OC                    | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; |  |                                      |
| OC                    | eurosid I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.        |  |                                      |
| NCBI_TaxID=3827;      |   |  |                                      |
| RN                    | [1]   |  |                                      |
| RP                    | SEQUENCE FROM N.A.  |  |                                      |
| RC                    | STRAN-ILC379; TISSUP-LEAF;  |  |                                      |
| RA                    | Ichinose Y., Tiemann K., Schwenger-Erger C., Toyoda K., Hein F.,      |  |                                      |
| RT                    | Hanselle T., Barz W.;   |  |                                      |
| RT                    | *Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and   |  |                                      |
| RT                    | Elicited Cell Cultures as Detected by Differential cDNA-              |  |                                      |
| RT                    | Hybridization.  |  |                                      |
| RL                    | Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.               |  |                                      |
| CC                    | -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.        |  |                                      |
| DR                    | EMBL; AB024994; BAA76422.1; -   |  |                                      |
| DR                    | HSSP; P01112; 1PLK.   |  |                                      |
| DR                    | InterPro; IPR003579; GTPase_Rab.                                      |  |                                      |
| DR                    | InterPro; IPR001806; Ras.transfrmng.                                  |  |                                      |
| DR                    | InterPro; IPR002078; Sig54_interact.                                  |  |                                      |
| DR                    | InterPro; IPR005225; Small_GTP.                                       |  |                                      |
| DR                    | Pfam; PF00071; ras; 1.  |  |                                      |
| DR                    | PRINTS; PR00449; RASTRNSFRMNG.  |  |                                      |
| DR                    | SMART; SM00175; RAB; 1.   |  |                                      |
| DR                    | TIGRFAMS; TIGR00231; small_GTP; 1.                                    |  |                                      |
| DR                    | PROSITE; PS00675; SIGNA54_INTERACT_1; UNKNOWN_1.                      |  |                                      |
| DR                    | GTP-binding; Lipoprotein.   |  |                                      |
| SK                    | SEQUENCE 202 AA; 22487 MW; 8CC4FCBAAC0C8CE CRC64;                     |  |                                      |
| Query Match           | 44.7%;  | Score 514.5;   | DB 10; Length 202;                   |
| Best Local Similarity | 47.3%;  | Pred. No. 1.7e-39;   |                                      |
| Matches               | 96;   | Conservative   | 47; Mismatches 53; Indels 7; Gaps 3; |
| QY                    | 23  | SPSYDLTGKVMVLGDTGVGKTCFLIQFDGAGLSGTFIATVGIDFRKNKVVTVDGVRVKLQ   | 82                                   |
| Db                    | 2   | NPWYDLFKLLIGDSGVKSCLLRFADDSYLD-SVISTIGVDFKIRTVEQDGKTIKLQ       | 60                                   |
| QY                    | 83  | IWDTAGQERPRSTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFTEIHEYAQRDVVIMILG  | 142                                  |
| Db                    | 61  | IWDTAGQERPRFTTSSYIRGAHGIYVYDVTQGESFNVNKQINLEIDRYASENVNKLVG     | 120                                  |
| QY                    | 143   | NKADMSRSVRSEDECTLAREYGVFFLETSAKTGMNVELAFATAKEIKYRAGHQ----      | 198                                  |
| Db                    | 121   | NKCDLAANKVVSSETAKAFADGIGPFMETYSAKNATNVEQAFMAAAEIKNRMASOPANN    | 180                                  |
| QY                    | 199   | ADEFSPQIRDYVESQKKRSSCS   | 221                                  |
| Db                    | 181   | ARPEVTQIRGQPNQK--SGCCS   | 201                                  |
| RESULT 6              |   |  |                                      |
| Q08153                |   |  |                                      |



```

Db      121 NKSCUTANKVWSEFAKAFADSMGIPFMETSNAKNATVEQAFNMAASIKDRMSQPAAA 180
      199 -ADEPSFOIRDYVESOKKRSSCCS 221
      181 NAREPTVQIRGQPVNOK--TSCCS 202

RESULT 10
Q40215
ID      Q40215      PRELIMINARY;      PRT;      216 AA.
AC      Q40215;
AL      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
RN      RAB8A.
GN      RAB8A.
OS      Lotus japonicus.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OC      NCBI_TaxID=34305;
RN      [1]
SE      SEQUENCE FROM N.A.
RP      TISSUE=ROOT NODULES;
RC      MEDLINE=972331679; PubMed=9076991;
RA      Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT      "Identification of new protein species among 33 different small GTP-
RT      binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT      of corresponding mRNAs in developing root nodules.";
RL      Plant J. 11:237-250(1997).
- 1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR      EMBL; J73944; CAA98172.1; -.
DR      HSP; P05713; 3RAB.
DR      InterPro; IPR003579; GTPase_Rab.
DR      InterPro; IPR001806; Ras_transfmg.
DR      InterPro; IPR002078; Sig54_interact.
DR      InterPro; IPR005225; Small_GTP.
DR      Pfam; PF00071; ras; 1.
DR      PRINTS; PR00449; RASTRNSFRMNG.
DR      SMART; SM00175; RAB; 1. small_GTP; 1.
DR      TIGRFAMs; TIGR00231; small_GTP; 1.
DR      PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW      GTP-binding; Lipoprotein.
SQ      SEQUENCE      216 AA; 23776 MW; 569926CCA8D1B489 CRC64;

Query Watch      44.3%; Score 509; DB 10; Length 216;
Best Local Similarity 45.2%; Pred. No. 6.1e-39;
Matches 103; Conservative 46; Mismatches 57; Indels 22; Gaps 5;

QY      1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIOFKDGAFLSGTFI 60
DB      1 MAGAAPAARAD-----YDIKLLLLGDSGVKSLLRSDGSFTT-SFI 46

QY      61 ATVGIDFRNKVTVYDGVVRVKLIQWDPAGQERFSVTHAYYRDAQALLLYLDITNKKSSFDN 120
DB      47 TTGIDFKRTIETLDDGKRIKLIQWDPAGQERFTITAYYRGAMGILLVYDVYDDEASFPN 106

QY      121 IRAWLTEIHEYAQDVVIMLGNKADM--SSERVISEDGETIAREYGVGFLETSAKTGMN 179
DB      107 IKNWIRNPIDQASDNVKNKTLVGNKADMDSKRAVPTSKQALADEYGIKFFETSAKTNLN 166

QY      180 VELAFIAIAKELKYR---AGHQADEPSFOIRDYVES---QKRSSCC 220
DB      167 VEEVFFSIARDIKQRLADTDSRAEPQTQIQNPDASASGGQAQKSSCC 214

RESULT 11
Q9SF91
ID      Q9SF91      PRELIMINARY;      PRT;      218 AA.
AC      Q9SF91;
AL      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-MAY-2000 (TREMBlrel. 13, Last sequence update)

```

Query Match 44.2%; Score 508.5; DB 5; Length 204;  
Best Local Similarity 49.8%; Pred. No. 6.3e-39;  
Matches 100; Conservative 42; Mismatches 54; Indels 5

```

Qy 25 SYDLTGKVMLLGDTGVCKTCLFQKDGAFSLGTFIATVIGIDFRNKVVTVVDCGVKQLQIW 84
Db 5 TYDLFLKLLIGDSGVGKTCILFRSDDAFTS-TFISTIGIDFKIKTVELRGKKIKLQIW 63
Qy 85 DTAGQFRSVTHAYYRDAQALLLLYDITNKSFDNRAMLTETHEVAQRDVVIMLGNK 144
Db 64 DTAGQFRSVTHAYYRDAQALLLLYDITNKSFDNRAMLTETHEVAQRDVVIMLGNK 123
Qy 145 ADMSSERVIRSEGETLAREYGVFFLETSARTGMNVELAFIAIAKE-LKYRAGHQADEPS 203
Db 124 CDWTDKRVNKEGEAREHIGIRFMETSAKSNINIERAFCELAAILDKTSGRESAENQ 183
Qy 204 FQIRDYVESQKR---SSCCS 221
Db 184 ERVIIDRRNQEKAPGYSKCCA 204

RESULT 13
Q40203 PRELIMINARY; PRT; 202 AA.
AC Q40203;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RAB1C.
GN RAB1C.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34309;
[1]
SEQUENCE FROM N.A.
RP TISSUE-ROOT NODES:
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
RT binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73932; CA98160.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 202 AA; 22633 MW; 91633EBFAD07662F CRC64;

Query Match 44.1%; Score 507.5; DB 10; Length 202;
Best Local Similarity 46.8%; Pred. No. 7.6e-39;
Matches 95; Conservative 46; Mismatches 55; Indels 7; Gaps 3;

Qy 23 SPVDLTKVMLLGDTGVCKTCLFQKDGAFSLGTFIATVIGIDFRNKVVTVVDCGVKQLQ 82
Db 2 NPEDYLFKLLIGDSGVGKSCLLIRADDSYLD-SYSTIGVDFKTVTEQDQKTKLQ 60
Qy 83 IWDTAGQFRSVTHAYYRDAQALLLLYDITNKSFDNRAMLTETHEVAQRDVVIMLGN 142
Db 61 IWDTAGQFRSVTHAYYRDAQALLLLYDITNKSFDNRAMLTETHEVAQRDVVIMLGN 120
Qy 143 NKADMSSERVIRSEGETLAREYGVFFLETSARTGMNVELAFIAIAKE-LKYRAGHQADEPS 198
Db 121 NKCDLTENKVVSYETAKAFADDEIGIPFMETSATNATNVEQAFMAAAEIKRNASQPVNN 180
Qy 199 ADEPSFQIRDYVESQKRSSCCS 221

```

```

Db 181 ARPETVOIRQPVNQK--SSCCS 201

RESULT 14
Q9LZD4 PRELIMINARY; PRT; 216 AA.
AC Q9LZD4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GTP-binding Protein-like (Putative GTP-binding protein).
GN F12E4_300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RP Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full Length cDNA of gene F12E4_300 (GI:7378637).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full Length cDNA of gene F12E4_300 (GI:7378637).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AL162751; CAB83313.1; -.
DR EMBL; AY056351; AAL07200.1; -.
DR EMBL; AY035124; AAK59629.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 24037 MW; 50C7D3AF4DC86F67 CRC64;

Query Match 44.1%; Score 507; DB 10; Length 216;
Best Local Similarity 47.5%; Pred. No. 9.4e-39;
Matches 103; Conservative 46; Mismatches 56; Indels 12; Gaps 4;

Qy 15 APEISPPCSPSYDLTKVMLLGDTGVCKTCLFQKDGAFSLGTFIATVIGIDFRNKVVTV 74

```

Db 4 APARA---RSDYDYLKLLIGDSGVKSCLLRFSDDFFTT-SFTTTIGIDFKIRTVEL 59

QY 75 DGVYKLIQWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFNIRAWLFEIHEYAQR 134

Db 60 DGKRIKLIQWDTAGQERFRITITAYYRGAMGILLVDVDTDESFNNIRNMKNIEQHASD 119

QY 135 DVVIMLLGNKADM-SSEVRISDEGETLAREYGVFPLETSAKTMNVELAFIAIAKELKY 193

Db 120 NVNKLIVGNKADMESKRAVPTAKGQALADEYGIKFFETSAKTNLNVENVFMSIARDIKQ 179

QY 194 R-----AGHQADEPSFIQIRDYVESQKRRSCCSFM 223

Db 180 RLUTEDTKRAEPOGKITRQDTAASSSTAESKACSYSV 216

RESULT 15

QBVWF9 PRELIMINARY; PRT; 216 AA.

AC Q8VWF9;

DT 01-MAR-2002 (TREMREL. 20, Created)

DT 01-MAR-2002 (TREMREL. 20, Last sequence update)

DT 01-JUN-2002 (TREMREL. 21, Last annotation update)

DE Ras-related protein RAB8-4 (Ras-related protein RAB8-2).

GN RAB8-4 OR RAB8-2.

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.

OX NCBI\_TaxID=4097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRIGHT YELLOW 2;

RA Torimoto N., Shimada K., Ito K., Yamamoto K.;

RT "Characterization of Rab8 from tobacco BY-2 cell.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB079023; BAB84325.1; -

DR EMBL; AB079021; BAB84323.1; -

DR InterPro; IPR003579; GTPase\_Rab.

DR InterPro; IPR001806; Ras\_trnsfmg.

DR InterPro; IPR002078; Sig54\_interact.

DR Pfam; PF00071; ras; 1.

DR PRINTS; PR00449; RASTRNSFRMNG.

DR SMART; SM00175; RAB; 1.

DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.

SQ SEQUENCE 216 AA; 23971 MW; D5E87FABD0D49E13 CRC64;

Query Match 44.1%; Score 507; DB 10; Length 216;

Best Local Similarity 46.9%; Pred. No. 9.4e-39;

Matches 100; Conservative 45; Mismatches 56; Indels 12; Gaps 4;

QY 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFDGAPLSGTFIATVGIIDFRNKVTVYDG 76

Db 3 APPARADYDLKLLIGDSGVKSCLLRFSDDFFTT-SFTTTIGIDFKIRTVELDG 61

QY 77 VVKLIQWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFNIRAWLFEIHEYAQRDV 136

Db 62 KRILQIWDTAGQERFRITITAYYRGAMGILLVDVDTDESFNNIRNMKNIEQHASD 121

QY 137 VIMLLGNKADM-SSEVRISDEGETLAREYGVFPLETSAKTMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMESKRAVPTSKGQALADEYGIKFFETSAKTNMNVVEVFMSIARDIKORL 181

QY 196 GHQADEPSFIQIRDYVESQK-----KRSCC 220

Db 182 SESDKTEPQAIRINQSDAGTSGQAQKSSCC 214

Search completed: June 18, 2003, 15:46:31

Job time : 87 secs



| Result No. | Score | Query |        | ID | Description         |
|------------|-------|-------|--------|----|---------------------|
|            |       | Match | Length |    |                     |
| 1          | 1150  | 100.0 | 223    | 1  | RB37_HUMAN          |
| 2          | 1081  | 94.0  | 223    | 1  | Q9Jkm7 mus musculus |
| 3          | 753   | 65.5  | 190    | 1  | RB26_MOUSE          |
| 4          | 745   | 64.8  | 190    | 1  | RB26_RAT            |
| 5          | 517.5 | 45.0  | 203    | 1  | SAS2_DICDI          |
| 6          | 516   | 44.9  | 206    | 1  | RAB8_MOUSE          |
| 7          | 513.5 | 44.7  | 207    | 1  | RAB8_HUMAN          |
| 8          | 512.5 | 44.6  | 215    | 1  | RAB1_BETUV          |
| 9          | 512   | 44.5  | 210    | 1  | RAB8_DISOM          |
| 10         | 507   | 44.1  | 258    | 1  | ARA5_ATH            |
| 11         | 502.5 | 43.7  | 207    | 1  | R88B_RAT            |
| 12         | 501   | 43.6  | 207    | 1  | R88B_HUMAN          |
| 13         | 500   | 43.5  | 202    | 1  | RIC1_ORYSA          |
| 14         | 498   | 43.3  | 208    | 1  | SAS1_DICDI          |
| 15         | 495   | 43.0  | 203    | 1  | YPT2_MAI2E          |
| 16         | 495   | 43.0  | 216    | 1  | ARA3_ATH            |
| 17         | 494   | 43.0  | 200    | 1  | RB10_CANFA          |
| 18         | 494   | 43.0  | 200    | 1  | RB10_HUMAN          |
| 19         | 490.5 | 42.7  | 200    | 1  | RAO1_DISOM          |
| 20         | 487   | 42.3  | 200    | 1  | YPT2_SCHPO          |
| 21         | 486   | 42.3  | 203    | 1  | YPT1_CHLRE          |
| 22         | 479.5 | 41.7  | 208    | 1  | YPT1_MAI2E          |
| 23         | 478   | 41.6  | 203    | 1  | YPT1_VOLCA          |
| 24         | 477.5 | 41.5  | 205    | 1  | RAB1_LYMST          |
| 25         | 476   | 41.4  | 217    | 1  | YPT2_VOLCA          |
| 26         | 473.5 | 41.2  | 203    | 1  | RB13_HUMAN          |
| 27         | 472   | 41.0  | 201    | 1  | RB1B_RAT            |
| 28         | 465.5 | 40.5  | 202    | 1  | RAB1_DISOM          |
| 29         | 464.5 | 40.4  | 203    | 1  | RV1L_YARLI          |
| 30         | 461   | 40.1  | 200    | 1  | RB10_RAT            |
| 31         | 459.5 | 40.0  | 203    | 1  | YPT1_NEUCR          |
| 32         | 458.5 | 39.9  | 205    | 1  | RB1A_HUMAN          |
| 33         | 456.5 | 39.7  | 205    | 1  | RB1A_RAT            |







```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56741; CAA40065.1; -
DR EMBL; S53268; AAB19681.1; -
DR EMBL; BC002977; AAH02977.1; -
DR EMBL; X56385; CAB56776.1; -
DR PIR; B36364; B36364.
DR PIR; B49647; B49647.
DR HSSP; P05713; 3RAB.
DR Genew; HGNC:7007; MEL.
DR MIM; 165040; -
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).
SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD036 CRC64;

Query Match 44.7%; Score 513.5; DB 1; Length 207;
Best Local Similarity 50.0%; Pred. No. 2.9e-38;
Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

QY 25 SYDTGKVMILGDTGVKTCFLIQKDGAFLSGFTFATVGDIFNKVYVVGVRVQLQIW 84
DB 4 TYDLFLKLLIGDGVGKTCFLFRSEDAF-NSTFISTIGIDFKIRTIKRIKLIQIW 62

QY 85 DTAGQERFRSVTHAYRDAQALLLYDITNKSSFNDINRAWLTEIHEYAQRDVVIMLGK 144
DB 63 DTAGQERFRITTAIRGAMGIMLYDITNKSEFNDINRWIRNIEEHSADVEKMLGNK 122

QY 145 ADMSSRVSSEGETLAREYGVFPLETSKATGMNVELAFIAIAELKYRAGHQAD---- 200
DB 123 CDVNDKQVSKERGEKALDYGIRKFMETSAKANINVENAFPTLARDIKAKMDKLEGNP 182

QY 201 EPSFQIRDYVESQKRSS 218
DB 183 QGSNGVKITPDQQRSS 200

RESULT 8
RAB1_BETVU
ID RAB1_BETVU STANDARD; PRT; 215 AA.
AC Q39433;
DT 15-JUL-1998 (Rel. 36, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein RAB1BV.
GN RAB1BV.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3355;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. D100 KS 38080;
RX MEDLINE=96307523; PubMed=8680950;

```

```

RA Dallery E., Quief S., Ben Jilany K.E., Kerckaert J.-P., Hagege D.;
RT 'Molecular cloning and structural analysis of cDNAs that encode 3
RL small GTP-binding proteins from sugar beet.';
CC C. R. Acad. Sci., III, Ser. Vie 319:91-97(1996).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49152; CAA89021.1; -
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT LIPID 212 212 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 215 AA; 23787 MW; 781CC5A17F29B8AA CRC64;

Query Match 44.6%; Score 512.5; DB 1; Length 215;
Best Local Similarity 47.6%; Pred. No. 3.7e-38;
Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

QY 19 SPR--CSPSYDLTGKVMILGDTGVKTCFLIQKDGAFLSGFTFATVGDIFNKVYVVDG 76
DB 3 APPARADYDLIRKLLIGDGVGKSCLLRFSDGSEFT-SFTTIGIDFKIRTIKRI 61

QY 77 VRVLQIWDTAGQERFRSVTHAYRDAQALLLYDITNKSSFNDINRAWLTEIHEYAQRDV 136
DB 62 KRILQIWDTAGQERFRITTAIRGAMGIMLYDITNKSEFNDINRWIRNIEEHSADSV 121

QY 137 VIMLGNKADM-SERVIRSEGETLAREYGVFPLETSKATGMNVELAFIAIAELKLYRA 195
DB 122 NKILVGNKADMESKRAVPTAKGALADEYGIKFEFSKATNLNVEYFFSIARDIKORL 181

QY 196 G----HQADEPSFOIRDYVESQK---RSSCC 220
DB 182 ADSDTQEAQPSITIKPADQSGNQAAKSACC 213

RESULT 9
RAB8_DISOM
ID RAB8_DISOM STANDARD; PRT; 210 AA.
AC P22128;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (ORA2).
OS Discopoge ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hymnosquales; Pristiogalea; Batoidae;
OC Torpediniformes; Narcinoidei; Narcinidae; Discopogae.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngsee J.K., Elferink L.A., Scheller R.H.;
RT "A family of ras-like GTP-binding proteins expressed in electromotor
RT neurons."
RL J. Biol. Chem. 266:2675-2680(1991).

```

```

CC  -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M38391; AAA49232.1; -
DR  PIR; B38625; B38625.
DR  HSSP; P05713; 3RAB.
DR  InterPro; IPR003579; GTPase_Rab.
DR  InterPro; IPR001230; Prenyl_site.
DR  InterPro; IPR001806; Ras_trnsmfmg.
DR  InterPro; IPR005225; Small_GTP.
DR  Pfam; PF00071; ras; 1.
DR  PRINTS; PR00449; RASTNSFRMNG.
DR  SMART; SM00175; RAB; 1.
DR  TIGRFAMs; TIGR00231; small_GTP; 1.
KW  GTP-binding; Prenylation; Lipoprotein.
FT  NP_BIND 15 22 GTP (BY SIMILARITY).
FT  NP_BIND 63 67 GTP (BY SIMILARITY).
FT  NP_BIND 121 124 GTP (BY SIMILARITY).
FT  DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT  LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
SQ  SEQUENCE 210 AA; 24164 MW; 0F73ED80DB89EEA CRC64;

Query Match 44.5%; Score 512; DB 1; Length 210;
Best Local Similarity 48.6%; Pred. No. 4e-38;
Matches 102; Conservative 44; Mismatches 50; Indels 14; Gaps 4;

QY 25 SYDLTGKVMLGDTGKTCFLIQKDAFLSGTFTATVGVDFRNKVVTDGVRVKLQIW 84
DB 4 TYDLKLLIGDSGVGKTCFLIFSEDAF-NTTFISTIGIDFKINTVELDGKIKLQIW 62

QY 85 DTAGQERFRSTHAYYRDAQALLLDITNKSFDFNIRAWLFEIHEYAQRDVVIMLGK 144
DB 63 DTAGQERFRFTTAYYRGAWGIMKYDITNKSFDFNIRAWLFEIHEYAQRDVVIMLGK 122

QY 145 ADMSSRVIRSEDETALREYGVPPLETSAKTMGMVELAFIAKELKYRAGHQADEPSF 204
DB 123 CDNNEKQVSKERGEKLAIDYGIKPLETSKSSINVEAFITLARDIMTKLNKKNENSL 182

QY 205 QIRDYVE-----SQKR--SSCCSPM 223
DB 183 Q--EAYDKLKSPKPKSOKKQLSFRCSLL 210

RESULT 10
ARA5_ARATH STANDARD; PRT; 258 AA.
AC P28188;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ARA-5.
GN ARA-5 OR ATIG02130 OR T7123.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,

```

```

RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.J., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.I., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
RA Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE OF 64-258 FROM N.A.
RC STRAIN=Various strains; TISSUE=Leaf;
RX MEDLINE=92084144; PubMed=1748311;
RA Anal T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
RA Matsui M.;
RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins
RT of Arabidopsis thaliana.";
RL Gene 108:259-264(1991).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; U89959; AAC24370.1; -
DR  EMBL; D01027; BAA00832.1; -
DR  PIR; PS0279; PS0279.
DR  HSSP; P05713; 3RAB.
DR  InterPro; IPR003579; GTPase_Rab.
DR  InterPro; IPR001806; Ras_trnsmfmg.
DR  InterPro; IPR005225; Small_GTP.
DR  Pfam; PF00071; ras; 1.
DR  PRINTS; PR00449; RASTNSFRMNG.
DR  SMART; SM00175; RAB; 1.
DR  TIGRFAMs; TIGR00231; small_GTP; 1.
KW  GTP-binding; Lipoprotein; Prenylation; Multigene family.
FT  NP_BIND 70 77 GTP (BY SIMILARITY).
FT  NP_BIND 118 122 GTP (BY SIMILARITY).
FT  NP_BIND 176 179 GTP (BY SIMILARITY).
FT  DOMAIN 92 100 EFFECTOR REGION (BY SIMILARITY).
FT  LIPID 255 255 GERANYL-GERANYL (BY SIMILARITY).
FT  LIPID 256 256 GERANYL-GERANYL (BY SIMILARITY).
SQ  SEQUENCE 258 AA; 29169 MW; 34D278A0C504638F CRC64;

Query Match 44.1%; Score 507; DB 1; Length 258;
Best Local Similarity 45.5%; Pred. No. 1.4e-37;
Matches 95; Conservative 50; Mismatches 56; Indels 8; Gaps 3;

QY 18 RSPPCSPSYDLTGKVMLGDTGKTCFLIQKDAFLSGTFTATVGVDFRNKVVTDGCV 77
DB 52 RSKTMPEYDYLKLLIGDSGVGKTCFLIFSEDAF-NTTFISTIGIDFKINTVELDGK 110

QY 78 RYKLOIWDTAGQERFRSTHAYYRDAQALLLDITNKSFDFNIRAWLFEIHEYAQRDV 137
DB 111 TIKLOIWDTAGQERFRFTTSSYRGANGIIIVDVTDEESFNQWQLSEIDRYASDNV 170

QY 138 IMLGNKADMSRVIRSEDETALREYGVPPLETSAKTMGMVELAFIAKELKYR--- 194
DB 171 KLVGNKSDLTENRAIPYETAKAFADGIPMETSAKDNTNVEQAFNMSASIKERMAS 230

QY 195 --AGHQADEPSFQIRDYVESQKRSSCCS 221
DB 231 QPAGNNARPPTVQIRGQPVQAK--NGCCS 257

```



```
DB 63 DTAGQERFRITTYAYRGAMGLVYDITNKSFNFNINRNINRIIEHASSDVERMILNK 122
QY 145 ADMSSERVRSDEGTAREYGVPELETSAKTMGMVELAFIAIAKELAYRAGHOA-DEPS 203
DB 123 CMNDKRVQSKERGLAIDYIKLETSKSSANVEAEFTLARDIMTKLNKRNKNDNS 182
QY 204 FOIRDYVESQKRSCCSF 222
DB 183 AGAGGPVKITENKSKTSF 201

RESULT 13
R1C1_ORYSA STANDARD; PRT; 202 AA.
ID SAS1_DICDI
AC P40392;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ras-related protein R1C1.
GN R1C1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-CV, Yamahoushi, TISSUE=Callus;
RX MEDLINE=94009718; PubMed=8405471;
RA Uchimiya H., Kidou S., Anai T., Umeda M., Aotsuka S., Tsuge T.,
RA Kato A.;
RT "Molecular structure of ras-related small GTP-binding protein genes
RT of rice plants and GTPase activities of gene products in Escherichia
RT coli.";
RL FEBS Lett. 332:282-286(1993).
CC -1- FUNCTION: POSSESS GTPASE ACTIVITY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S66160; AAB28535.1; -
DR PIR; S38740; S38740.
DR HSP; P01112; IPLK.
DR InterPro; IPR003579; GTPase.Rab.
DR InterPro; IPR001806; Ras.transfmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 200 200 PALMITATE (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 202 AA; 22561 MW; 5A1CB5F93256B659 CRC64;

Query Match 43.5%; Score 500; DB 1; Length 202;
Best Local Similarity 46.3%; Pred. No. 4.4e-37;
Matches 94; Conservative 49; Mismatches 54; Indels 6; Gaps 3;

QY 23 SPVDLTGKVLGDTGCVGKTCFLQKDGAFSLGTFIATVGTIDFNKVVTVGVRVKLQ 82
DB 2 NPEYDLFLKLLIGDSGVKSCLLRFAADDSYLE-SYISTIGVDFKIRTVEQDGKTIKQ 60
```

```
QY 83 IWDTAGQERFRSVTHAYYRQAQALLLYDITNKSFNFNINRNINRIIEHAYQRDVVMILG 142
DB 61 IWDTAGQERFRITTSYIRGAHGIIVYDVTQDSFNNKQWLNEDRIDRYASEYNKILVG 120
QY 143 NKADSSERVRSDEGTAREYGVPELETSAKTMGMVELAFIAIAKELAYRAGHOA-DEPS 198
DB 121 NKCDLAENRVVSYEAGKALADEIGIPELETSAKDATNVERKAFMTMAGEIKRNRMSQGRTN 180
QY 199 ADEPSFOIRDYVESQKRSCCS 221
DB 181 ASKPA-TVOMPROQVPAQSSCCS 202

RESULT 14
SAS1_DICDI STANDARD; PRT; 208 AA.
ID SAS1_DICDI
AC P20790;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GTP-binding protein SAS1.
GN SASA OR SAS1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220623; PubMed=2109188;
RA Saxe S.A., Kimmel A.R.;
RT "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
RT with sequence similarities to essential genes in Saccharomyces
RT cerevisiae.";
RL Mol. Cell. Biol. 10:2367-2378(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89209367; PubMed=3149563;
RA Saxe S.A., Kimmel A.R.;
RT "Genes encoding novel GTP-binding proteins in Dictyostelium.";
RL Dev. Genet. 9:259-265(1988).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34457; AAA33249.1; -
DR PIR; A34716; A34716.
DR PIR; A61571; A61571.
DR HSP; P05713; 3RAB.
DR DictyDb; DD05020; sasA.
DR InterPro; IPR003579; GTPase.Rab.
DR InterPro; IPR001806; Ras.transfmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR GTP-binding; Lipoprotein; Palmitate.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT LIPID 207 207 PALMITATE (BY SIMILARITY).
FT LIPID 208 208 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 208 AA; 23243 MW; BF681EF5C910019E CRC64;

Query Match 43.3%; Score 498; DB 1; Length 208;
Best Local Similarity 46.6%; Pred. No. 6.8e-37;
```



| Matches       |     | 95; Conservative   | 49; Mismatches | 50; Indels | 10; Gaps | 3; |
|---------------|-----|--|----------------|------------|----------|----|
| QY            | 23  | SPSYDLTGKVMLLGGTGVGKTCFLIQKDGAFLSGTFIATVGIDFRKNVVDGVRVKLQ    | 82             |            |          |    |
| Ddb           | 9   | SAAAYDYLKLLIGDSGVGSKCLLRFSEDSF-TPSFITTIGIDFKIRTIELSGKRILQ    | 67             |            |          |    |
| QY            | 83  | IWDTAGOQERFSRVTTHAAYRDQAQLLLDYITNKSSFDNIRAWKTETHEVAQRDVVMLLG | 142            |            |          |    |
| Ddb           | 68  | IWDTAGOQERFTTTAYYRGAMGILLVYDVTDEKSFGNRINRWTRNEQHATDSVNKMILG  | 127            |            |          |    |
| QY            | 143 | NKADMSRSRVASEDCETLAREYGVPVFLETSARTGMNVLAFLAIKAELKYRAGHOADEP  | 202            |            |          |    |
| Ddb           | 128 | NKCDMAEKKVDDSRGSGLADEYGIKFLETSAKNSINVVEAFISLANDIKKR---MIDTP  | 184            |            |          |    |
| QY            | 203 | SFOIR-----DYVESQKKRSSCC                                      | 220            |            |          |    |
| Ddb           | 185 | NEQPQVVQPGTNLGANNKKACC                                       | 208            |            |          |    |
| RESULT 15     |     |  | PRT; 203 AA.   |            |          |    |
| YPT2_MAIZE    |     |  |                |            |          |    |
| ID YPT2_MAIZE |     | STANDARD;  |                |            |          |    |

```

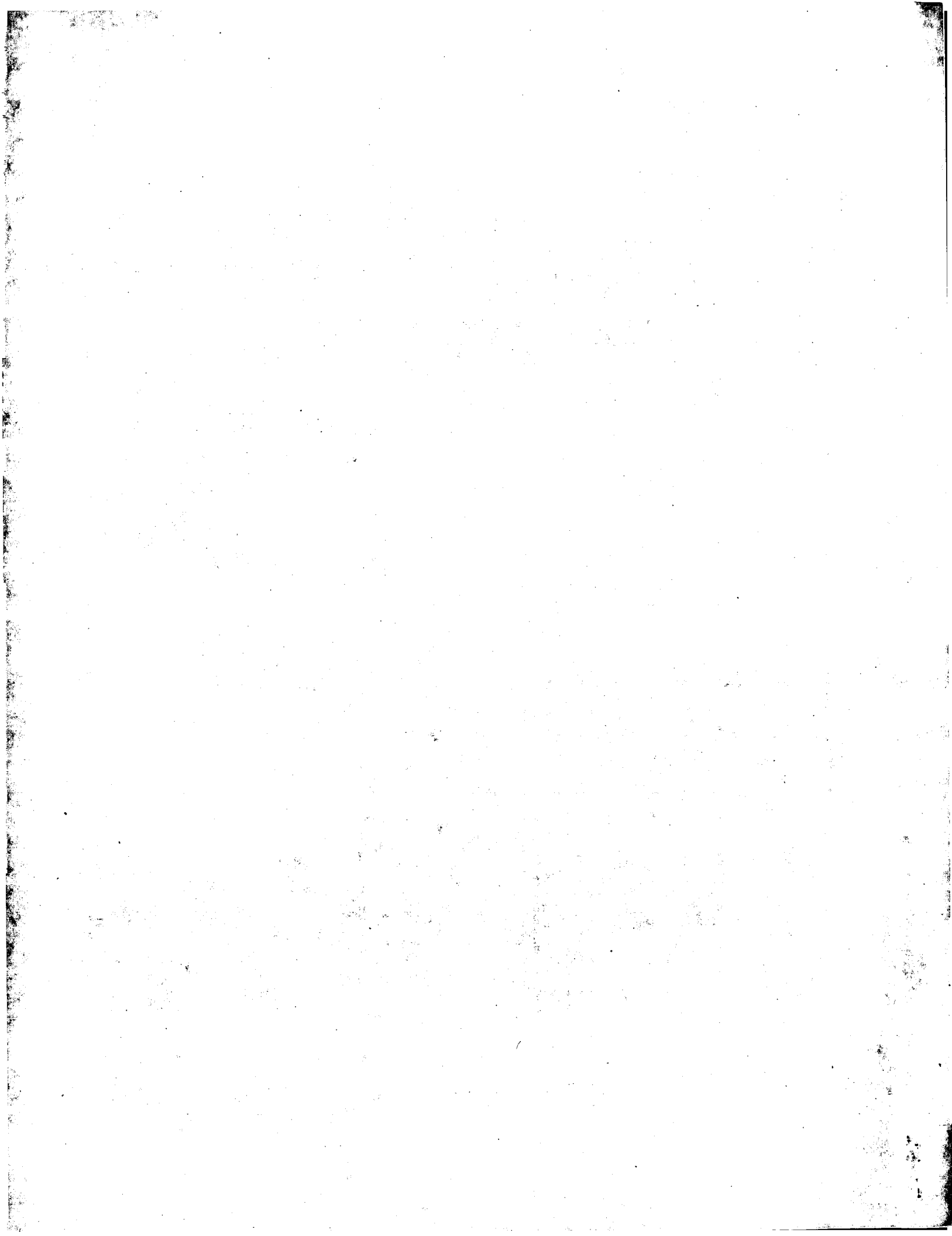
FT 200 PALMITATE (BY SIMILARITY).
LIPID 201 GERANYL-GERANYL (BY SIMILARITY).
LIPID 202
SQ SEQUENCE 203 AA; 22475 MW; E241326E7ACD1B8A CRC64;

Query Match 43.0%; Score 495; DB 1; Length 203;
Best Local Similarity 45.1%; Pred. No. 1.2e-36;
Matches 92; Conservative 51; Mismatches 53; Indels 8; Gaps 3;

QY 23 SP$YDLTGKVMLLGGTGVKTCFLQKDGAFSLGTFIATVGIDFRNKVVTVDGVKVLQ 82
Db 2 NPEYDYLKLLIGDSGVGSKLLRFADDSYLD-SYISTIGVDFIRIVEQDGKTKLQ 60
QY 83 IWTAGQERFRTVTHAYRYDAQALLLYDITKNKSPDNTRAMLTETHEYAQRDVIMLIG 142
Db 61 IWTAGQERFRTISYSGAHGIIIVYDVTQESFNKQWLNEIDRYASDNVNKLLVG 120
QY 143 NKADMSSEVIRSEGETLAREGVGPFLETSAKTGMNVELAFIALEKLYRAGHQ --- 198
Db 121 NKSDUTANKVATETAKAFADENGIPFMTSARNATNVQOAFMAAASIKDRMASQAAA 180
QY 199 -ADEPSFQIRDYVESOKRRSSCCS 221
Db 181 NARPATVQIRGQPVNQK--TSCCS 202

Search completed: June 18, 2003, 15:48:29
Job time : 33 secs

```



GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:37:02 ; Search time 71 Seconds  
 (without alignments)  
 418.520 Million cell updates/sec

Title: US-09-817-199B-2

Perfect score: 1150

Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDVYESOKKSSCCSEFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1          | 1150  | 100.0         | 232    | 22 | AA41696     |
| 2          | 1144  | 99.5          | 226    | 22 | AAU17119    |
| 3          | 1133  | 98.5          | 222    | 22 | AAU17541    |
| 4          | 977   | 85.0          | 191    | 22 | AAU17541    |
| 5          | 977   | 85.0          | 191    | 22 | AAU17541    |
| 6          | 977   | 85.0          | 191    | 22 | AAU17541    |
| 7          | 969   | 84.3          | 191    | 23 | AAU17541    |
| 8          | 946   | 82.3          | 191    | 23 | AAU17541    |
| 9          | 940   | 81.7          | 191    | 23 | AAU17541    |
| 10         | 926   | 80.5          | 191    | 20 | AAU17541    |

|    |       |      |     |    |          |
|----|-------|------|-----|----|----------|
| 11 | 770   | 67.0 | 158 | 22 | ABG17244 |
| 12 | 697   | 60.6 | 666 | 22 | ABG17578 |
| 13 | 666   | 57.9 | 190 | 19 | AAW80747 |
| 14 | 535   | 46.5 | 139 | 22 | AAU17123 |
| 15 | 514.5 | 44.7 | 221 | 23 | ABP41333 |
| 16 | 513.5 | 44.7 | 213 | 22 | ABU11916 |
| 17 | 509   | 44.3 | 218 | 21 | ABG19220 |
| 18 | 508.5 | 44.2 | 204 | 22 | ABG19670 |
| 19 | 507   | 44.1 | 253 | 21 | AAU53944 |
| 20 | 507   | 44.1 | 253 | 21 | AAU53944 |
| 21 | 507   | 44.1 | 258 | 21 | AAU53944 |
| 22 | 505.5 | 44.0 | 215 | 21 | AAU53944 |
| 23 | 502   | 43.7 | 216 | 21 | AAU53944 |
| 24 | 501   | 43.6 | 203 | 21 | AAU53944 |
| 25 | 501   | 43.6 | 207 | 22 | AAU53944 |
| 26 | 501   | 43.6 | 207 | 22 | AAU53944 |
| 27 | 500   | 43.5 | 216 | 21 | AAU53944 |
| 28 | 500   | 43.5 | 216 | 21 | AAU53944 |
| 29 | 498   | 43.3 | 218 | 21 | AAU53944 |
| 30 | 498   | 43.3 | 218 | 21 | AAU53944 |
| 31 | 497.5 | 43.3 | 202 | 21 | AAU53944 |
| 32 | 496.5 | 43.2 | 202 | 21 | AAU53944 |
| 33 | 494   | 43.0 | 200 | 21 | AAU53944 |
| 34 | 494   | 43.0 | 200 | 21 | AAU53944 |
| 35 | 494   | 43.0 | 200 | 21 | AAU53944 |
| 36 | 488   | 42.4 | 224 | 21 | AAU53944 |
| 37 | 488   | 42.4 | 234 | 21 | AAU53944 |
| 38 | 487.5 | 42.4 | 201 | 21 | AAU53944 |
| 39 | 487.5 | 42.4 | 207 | 22 | AAU53944 |
| 40 | 487   | 42.3 | 209 | 22 | AAU53944 |
| 41 | 483   | 42.0 | 218 | 21 | AAU53944 |
| 42 | 483   | 42.0 | 221 | 21 | AAU53944 |
| 43 | 483   | 42.0 | 246 | 21 | AAU53944 |
| 44 | 481.5 | 41.9 | 199 | 21 | AAU53944 |
| 45 | 481.5 | 41.9 | 199 | 21 | AAU53944 |

# ALIGNMENTS

RESULT 1  
 AA41696  
 ID AA41696 standard; Protein; 232 AA.  
 XX  
 AC AA41696;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 6627.  
 XX  
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.

Novel human diagno  
 Drosophila melanog  
 Human RAB protein,  
 Novel signal trans  
 Human ovarian anti  
 Human rab8 homolog  
 Arabidopsis thalia  
 Drosophila melanog  
 Arabidopsis thalia  
 Arabidopsis thalia  
 Arabidopsis thalia  
 Zea mays protein f  
 Arabidopsis thalia  
 Arabidopsis thalia  
 Amino acid sequenc  
 Human protein sequ  
 Arabidopsis thalia  
 Arabidopsis thalia  
 Amino acid sequenc  
 Human RAB10 protei  
 Human RAB10 protei  
 Human RAB10 protei

XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
DR N-PSDB; AA160852.  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX Example 2; SEQ ID NO 627; 10078pp; English.  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA136642-AA142213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX Sequence 232 AA;  
SQ

Query Match 100.0%; Score 1150; DB 22; Length 232;  
Best Local Similarity 100.0%; Pred. NO. 9.6e-120;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGTCGAVATGCEAPERSPPSPVDLTGKVMLLGDTGKTCFLIOFKDCAFLSGRFI 60  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
10 MTGTCGAVATGCEAPERSPPSPVDLTGKVMLLGDTGKTCFLIOFKDCAFLSGRFI 69  
QY 61 ATGVIDFRNKVTVGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
70 ATGVIDFRNKVTVGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 129  
QY 121 IRAWLTIHEYAQRDVVIMLLGNKADMSERVIRSEDGETLAREYGVPPLETSAKTGMV 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
130 IRAWLTIHEYAQRDVVIMLLGNKADMSERVIRSEDGETLAREYGVPPLETSAKTGMV 189  
QY 181 ELAFLAIKELKVRAGHQADEFSQIRDYVESQKRSSCCSFM 223  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
190 ELAFLAIKELKVRAGHQADEFSQIRDYVESQKRSSCCSFM 232

RESULT 2  
AAU17119  
ID AAU17119 standard; Protein; 226 AA.  
XX  
AC AAU17119;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Novel signal transduction pathway protein, Seq ID 684.  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.  
XX Homo sapiens.  
OS WO200154733-A1.  
PN  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01312.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249284.  
 PR 17-NOV-2000; 2000US-0249285.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254037.

PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-465460/50.  
 DR N-PSDB; AAS27036.  
 XX  
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders  
 XX  
 PS Claim 1; SEQ ID No 684; 880pp; English.  
 XX  
 CC The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
 CC pathway protein, amino acid sequences of the invention.  
 XX

Query Match 99.5%; Score 1144; DB 22; Length 226;  
 Best Local Similarity 99.6%; Pred. No. 4.3e-119;  
 Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLTFQKDFLSTGTFI 60  
 Db 4 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLTFQKDFLSTGTFI 63  
 Qy 61 ATVGIDFRNKVVTVGVRVKLQIWDTAGOERFRSVTHAYYRDQAALLLYDITNKKSSFDN 120  
 Db 64 ATVGIDFRNKVVTVGVRVKLQIWDTAGOERFRSVTHAYYRDQAALLLYDITNKKSSFDN 123  
 Qy 121 IRAWLTEIHEYAQRDVVIMLLGNKADMSRVSERSEGETLAREYGVFPFLETSARTGMNV 180  
 Db 124 IRAWLTEIHEYAQRDVVIMLLGNKADMSRVSERSEGETLAREYGVFPFLETSARTGMNV 183  
 Qy 181 ELAFLATAKELKYRAGHQADEPSPQIRDYVESQKRKSSCCSFM 223  
 Db 184 ELAFLATAKELKYRAGHQADEPSPQIRDYVESQKRKSSCCSFM 226

RESULT 3  
 AAU17541  
 ID AAU17541 standard; Protein; 222 AA.  
 XX  
 AC AAU17541;  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Novel signal transduction pathway protein, Seq ID 1106.  
 XX  
 KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;

KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX WO200154733-A1.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01312.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251989.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-465460/50.  
 DR N-PSDB; RAS27458.  
 DR  
 XX  
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders -  
 XX  
 XX  
 PS Claim 1: SEQ ID No 1106; 880pp; English.  
 XX  
 XX The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
 CC pathway protein, amino acid sequences of the invention.  
 XX  
 XX  
 XX Query Match 98.5%; Score 1133; DB 22; Length 222;  
 XX Best Local Similarity 99.1%; Pred. No. 7.1e-118;  
 XX Matches 220; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 TGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGKTCFLIQKDGAFLSGTFIA 61  
 DB 1 TGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGKTCFLIQKDGAFLSGTFIA 60  
 QY 62 TVGIDFNKVVTVGVRVKLIQWDTAGQERPSVTHAYYRDAQALLLYDITNKSSFDNI 121  
 DB 61 TVGIDFNKVVTVGVRVKLIQWDTAGQERPSVTHAYYRDAQALLLYDITNKSSFDNI 120  
 QY 122 RAWLTETHEYAQRDVVIMLLGNKADMSERVIRSEDGETFLAREYGVPPLETSAKTMNVE 181  
 DB 121 RAWLTETHEYAQRDVVIMLLGNKADMSERVIRSEDGETFLAREYGVPPLETSAKTMNVE 180  
 QY 182 LAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSPM 223  
 DB 181 LAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSPM 222  
 RESULT 4  
 AAG67156  
 ID AAG67156 standard; Protein; 191 AA.  
 XX  
 XX AAG67156;  
 XX  
 XX 13-NOV-2001 (first entry)  
 XX

DE Amino acid sequence of human 32712 G-protein.  
 XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;  
 KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;  
 KW adult respiratory distress syndrome; Goodpasture's syndrome;  
 KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;  
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;  
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;  
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;  
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;  
 KW chronic bacterial meningococcal meningitis; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; stroke; Huntington's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164887-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 XX 27-FEB-2001; 2001WO-US06292.  
 PF  
 XX 29-FEB-2000; 2000US-0185606.  
 PR  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX Meyers RA;  
 PI  
 XX WPI; 2001-550182/61.  
 DR N-PSDB; AAH75184.  
 DR  
 XX Novel human small G-protein polypeptides and polynucleotides for  
 PT treating lung disorders, liver disorders and brain disorders -  
 XX  
 XX Claim 8; Fig 26; 151pp; English.  
 XX  
 XX The present sequence represents a human G-protein. The specification  
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The  
 CC G-protein polypeptides and polynucleotides are useful as a target for  
 CC diagnosis and treatment of G-protein mediated or related disorders,  
 CC and for identifying agonists and antagonists for diagnosis and  
 CC treatment. They are useful for treating disorders of lung (e.g.  
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory  
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,  
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver  
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,  
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and  
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute  
 CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic  
 CC bacterial meningococcal meningitis, multiple sclerosis, amyotrophic lateral  
 CC sclerosis, stroke and Huntington's disease).  
 XX  
 XX Sequence 191 AA;  
 QY Query Match 85.0%; Score 977; DB 22; Length 191;  
 QY Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
 QY Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 MLLSDGTGKTCFLIQKDGAFLSGTFIATVGIDFNKVVTVGVRVKLIQWDTAGQERF 92  
 DB 1 MLLSDGTGKTCFLIQKDGAFLSGTFIATVGIDFNKVVTVGVRVKLIQWDTAGQERF 60  
 QY 93 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTETHEYAQRDVVIMLLGNKADMSERV 152  
 DB 61 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTETHEYAQRDVVIMLLGNKADMSERV 120  
 QY 153 IRSJDEGTFLAREYGVPPLETSAKTMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 212  
 DB 121 IRSJDEGTFLAREYGVPPLETSAKTMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 180  
 QY 213 QKRSSCCSPM 223  
 DB 181 QKRSSCCSPM 191





CC protein 3 (RASP-3). RASP-3 shares 72% sequence identity with  
 CC Rab26. It includes regions that resemble the signature sequences  
 CC of the GTP-binding Ras superfamily, and family signatures of Ran,  
 CC a subfamily of Ras proteins which function in nucleocytoplasmic  
 CC transport, RNA synthesis, processing and export, and cell cycle  
 CC checkpoint control, as well as family signatures of transforming  
 CC protein p21, a subfamily of GTP-binding Ras proteins which have  
 CC been implicated in a number of tumours. Northern analysis showed  
 CC expression of RASP-3 in haematopoietic and immunological cDNA  
 CC libraries, all of which were associated with inflammation and the  
 CC immune response. The invention provides 7 novel human Ras proteins  
 CC (RASP-1 to -7) and polynucleotides, expression vectors, host cells,  
 CC antibodies, agonists and antagonists. It also provides methods for  
 CC diagnosing, treating or preventing disorders associated with RASP  
 CC expression, especially cancer and immune disorders.

XX SQ Sequence 191 AA;

Query Match 85.0%; Score 977; DB 23; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWDTAGQERF 92  
 DB 1 MLLGDTGVGTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWDTAGQERF 60

QY 93 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIYAAQRDVVIMLLGNKADMSSERV 152  
 DB 61 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIYAAQRDVVIMLLGNKADMSSERV 120

QY 153 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 212  
 DB 121 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 180

QY 213 QKKRSSCCSFM 223  
 DB 181 QKKRSSCCSFM 191

RESULT 7  
 AAM52190  
 ID AAM52190 standard; Protein: 191 AA.  
 AC AAM52190;  
 XX 07-FEB-2002 (first entry)  
 DT Human degranulation regulator SEQ ID NO 3.  
 DE Human degranulation regulator  
 XX Degranulation; mast cell; human; mouse; antiallergic.  
 KW Homo sapiens.  
 XX WO200179478-A1.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-JP03268.  
 XX 19-APR-2000; 2000JP-0118408.  
 XX (DAIN ) DAINIPPON PHARM CO LTD.  
 XX Yamada T, Ido M;  
 PI WPI: 2002-041335/05.  
 DR N-PSDB: ABA02774.  
 XX Mast cell degranulation controller for treatment of allergies -  
 XX Claim 11; Page 64-65; 85pp; Japanese.  
 XX The invention relates to a protein for regulating degranulation of mast

CC cells (degranulation regulators) and the encoding polynucleotides, with  
 CC antiallergic activity, used in the treatment of allergies associated  
 CC with degranulation of mast cells.

XX SQ Sequence 191 AA;

Query Match 84.3%; Score 969; DB 23; Length 191;  
 Best Local Similarity 99.0%; Pred. No. 1.1e-99;  
 Matches 189; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWDTAGQERF 92  
 DB 1 MLLGDSGVGTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVGVRVKLQIWDTAGQERF 60

QY 93 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIYAAQRDVVIMLLGNKADMSSERV 152  
 DB 61 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIYAAQRDVVIMLLGNKADMSSERV 120

QY 153 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 212  
 DB 121 IRSEGETLAREYGVPPLETSAKTGMNVELAFLAIKELKYRAGHQADEPSFQIRDYVES 180

QY 213 QKKRSSCCSFM 223  
 DB 181 QKKRSSCCSFM 191

RESULT 8  
 AAM52189  
 ID AAM52189 standard; Protein: 191 AA.  
 AC AAM52189;  
 XX 07-FEB-2002 (first entry)  
 DT Mouse degranulation regulator SEQ ID NO 2.  
 DE Mouse degranulation regulator  
 XX Degranulation; mast cell; human; mouse; antiallergic.  
 KW Mus musculus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 83  
 FT /note= "Encoded by CAG"  
 XX WO200179478-A1.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-JP03268.  
 XX 19-APR-2000; 2000JP-0118408.  
 XX (DAIN ) DAINIPPON PHARM CO LTD.  
 XX Yamada T, Ido M;  
 PI WPI: 2002-041335/05.  
 DR N-PSDB: ABA02773.  
 XX Mast cell degranulation controller for treatment of allergies -  
 XX Claim 10; Page 63-64; 85pp; Japanese.  
 XX The invention relates to a protein for regulating degranulation of mast  
 CC cells (degranulation regulators) and the encoding polynucleotides, with  
 CC antiallergic activity, used in the treatment of allergies associated  
 CC with degranulation of mast cells.

XX SQ Sequence 191 AA;

Query Match 82.3%; Score 946; DB 23; Length 191;  
 Best Local Similarity 96.3%; Pred. No. 4.2e-97;

| Matches | 184; | Conservative   | 2;  | Mismatches | 5; | Indels | 0; | Gaps | 0; |
|---------|------|--|-----|------------|----|--------|----|------|----|
| QY      | 33   | MLLGDSGVGKTCFLIQFKDGAFLSGTFTIATVGIDFRNKVTVDGRVKQLQIWDTAGQERF   | 92  |            |    |        |    |      |    |
| Db      | 1    | MLLGDSGVGKTCFLIQFKDGAFLSGTFTIATVGIDFRNKVTVDGRVKQLQIWDTAGQERF   | 60  |            |    |        |    |      |    |
| QY      | 93   | RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIETHEYAQRDQVVMILGNKADMSERV   | 152 |            |    |        |    |      |    |
| Db      | 61   | RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIETHEYAQRDQVVMILGNKADMSERV   | 120 |            |    |        |    |      |    |
| QY      | 153  | IRSEDEGTALAREGVGFPLETSAGTKGMNVELAFLAIKELKYRAGHQADEPSFQIRDVYES  | 212 |            |    |        |    |      |    |
| Db      | 121  | IRSEDEGTALAREGVGFPEXETSAKTKGMNVELAFLAIKELKYRAGHQADEPSFQIRDVYES | 180 |            |    |        |    |      |    |
| QY      | 213  | QKRRSCCSPM   | 223 |            |    |        |    |      |    |
| Db      | 181  | QKRRSCCSEF   | 191 |            |    |        |    |      |    |

|           |   |
|-----------|---|
| RESULT 10 |   |
| AAV30132  |   |
| ID        | AAV30132 standard; Protein; 191 AA.                                     |
| XX        |   |
| AC        | AAV30132;   |
| XX        |   |
| XX        | 27-OCT-1999 (first entry)   |
| DT        |   |
| XX        |   |
| XX        | An exocytotic protein designated Exo2.                                  |
| XX        |   |
| KW        | Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;         |
| DE        | Inflammation.   |
| KW        |   |
| XX        |   |
| XX        | Mus sp.   |
| OS        |   |
| XX        |   |
| XX        |   |
| PH        | Key Location/Qualifiers   |
| FT        | Misc-difference 5   |
| FT        | /note= "encoded by GAN"   |
| FT        | Misc-difference 27  |
| FT        | /note= "encoded by TTN"   |
| FT        |   |
| XX        |   |
| XX        | WO9942586-A2.   |
| PN        |   |
| XX        |   |
| PD        | 26-AUG-1999.  |
| XX        |   |
| XX        | 23-FEB-1999; 99WO-US03909.  |
| PF        |   |
| XX        |   |
| XX        | 26-MAY-1998; 98US-0086650.  |
| PR        |   |
| PR        | 23-FEB-1998; 98US-0075534.  |
| XX        |   |
| XX        | (RIGE-) RIGEL PHARM INC.  |
| PA        |   |
| XX        |   |
| XX        |   |
| PI        | Anderson D, Fisher J, Huang B, Lorens J, Luo Y;                         |
| PI        | Shen M;   |
| XX        |   |
| DR        | WPI; 1999-518605/43.  |
| DR        | N-PSDE; AAX86720.   |
| XX        |   |
| XX        |   |
| PT        | New exocytotic proteins useful for diagnosis and treatment of           |
| PT        | exocytosis-mediated conditions and in drug screening                    |
| XX        |   |
| PS        | Claim 12; Fig 7; 53pp; English.   |
| XX        |   |
| XX        |   |
| CC        | The present sequence represents an exocytotic protein designated Exo2.  |
| CC        | The protein has an inhibitory effect on exocytosis (i.e. the fusion of  |
| CC        | secretory vesicles with the cellular plasma membrane) and can be        |
| CC        | administered therapeutically to treat or prevent exocytosis-mediated    |
| CC        | disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving   |
| CC        | undesirable release of compounds via exocytosis e.g. inflammation       |
| CC        | mediated by the release compounds such as histamine. Exo2 is useful     |
| CC        | diagnostically and to produce antibodies useful to purify the proteins  |
| CC        | and therapeutically to reduce or eliminate the biological activity of   |
| CC        | the protein. Exo2 nucleic acids can be used therapeutically to increase |
| CC        | Exo2 activity in cells by known gene therapy techniques. They can also  |
| CC        |   |

CC be used to produce probes or primers to isolate Exo2 proteins from other  
CC organisms, especially humans. The nucleic acids, host cells and proteins  
CC are useful in screening assays to identify binding agents, especially  
CC drug screening assays to identify agonists and antagonists useful  
CC therapeutically to enhance or reduce Exo2 activity.  
XX  
SQ Sequence 191 AA;  
  
Query Match 80.5%; Score 926; DB 20; Length 191;  
Best Local Similarity 94.8%; Pred. No. 7.2e-95;  
Matches 181; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
QY 33 MLIGDTGCGTCTLIQFKDCAFLSGTFIATVGIDFRNKVTVGVVRVKLIQWDTAGQERF 92  
DB 1 MLIGXSGVGTCCTLIQFKDCAFLSGTFIATVGIDFRNKVTVGVVRVKLIQWDTAGQERF 60  
  
QY 93 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLLGNKADMSERV 152  
DB 61 RSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHIHEYAQRDVVIMLLGNKADMSERV 120  
  
QY 153 IRSEGETLAREYGVFPLETSAKTGMNVELAFIAIAKELKYRAGHQADEPSFOIRDYVES 212  
DB 121 IRSEGETLAREYGVFPLETSAKTGMNVELAFIAIAKELKYRAGROPDEPSFOIRDYVES 180  
  
QY 213 QKKRSSCCSPM 223  
DB 181 QKKRSSCCSEV 191  
  
RESULT 11  
ABG17244  
ID ABG17244 standard; Protein; 158 AA.  
AC ABG17244;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #17235.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX N-PSDB; AAS81431.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity -  
XX  
XX Claim 20; SEQ ID NO 47603; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX diagnostic chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC the polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 158 AA;  
  
Query Match 67.0%; Score 770; DB 22; Length 158;  
Best Local Similarity 99.3%; Pred. No. 1.4e-77;  
Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60  
DB 10 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 69  
  
QY 61 ATVGIDFRNKVTVGVVRVKLIQWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 120  
DB 70 ATVGIDFRNKVTVGVVRVKLIQWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 129  
  
QY 121 IRAWLTHIHEYAQRDVVIMLLGNKADMS 149  
DB 130 IRAWLTHIHEYAQRDVVIMLLGNKADMS 158  
  
RESULT 12  
ABB71578  
ID ABB71578 standard; Protein; 566 AA.  
XX  
XX ABB71578;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 41526.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
XX  
XX N-PSDB; ABL15681.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
XX Disclosure; SEQ ID NO 41536; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent



PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184564.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225477.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 18-AUG-2000; 2000US-0225759.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236357.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465460/50.  
 DR N-PSDB; AAS27040.  
 XX  
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:44:07 ; Search time 40 Seconds  
(without alignments)  
535.949 Million cell updates/sec

Title: US-09-817-199b-2  
Perfect score: 1150  
Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESOKKRSCCSFEM 223  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Query Match % | Score | Length | ID | Description |
|------------|---------------|-------|--------|----|-------------|
| 1          | 742           | 64.5  | 190    | 2  | JC2528      |
| 2          | 533           | 46.3  | 196    | 2  | Ti5123      |
| 3          | 517.5         | 45.0  | 203    | 2  | B34716      |
| 4          | 516           | 44.9  | 206    | 2  | I78851      |
| 5          | 513.5         | 44.7  | 207    | 2  | B49647      |
| 6          | 513.5         | 44.7  | 207    | 2  | B36364      |
| 7          | 512.5         | 44.6  | 215    | 2  | T14565      |
| 8          | 507           | 44.1  | 216    | 2  | T48378      |
| 9          | 507           | 44.1  | 258    | 2  | B86153      |
| 10         | 506           | 44.0  | 216    | 2  | S33900      |
| 11         | 505.5         | 44.0  | 201    | 2  | T28971      |
| 12         | 500           | 43.5  | 202    | 2  | S38740      |
| 13         | 500           | 43.5  | 216    | 2  | T45901      |
| 14         | 498           | 43.3  | 208    | 2  | A34716      |
| 15         | 497.5         | 43.3  | 202    | 2  | S41430      |
| 16         | 497.5         | 43.3  | 209    | 2  | B38625      |
| 17         | 497.5         | 43.3  | 215    | 2  | S57478      |
| 18         | 495           | 43.0  | 203    | 2  | B38202      |
| 19         | 495           | 43.0  | 216    | 2  | S57471      |
| 20         | 495           | 43.0  | 216    | 2  | JS0640      |
| 21         | 494           | 43.0  | 200    | 2  | D36364      |
| 22         | 490.5         | 42.7  | 200    | 2  | A38625      |
| 23         | 490.5         | 42.7  | 215    | 2  | S57462      |
| 24         | 489           | 42.5  | 222    | 2  | T14405      |
| 25         | 488           | 42.4  | 203    | 2  | S34253      |
| 26         | 487           | 42.3  | 200    | 2  | S12790      |
| 27         | 486           | 42.3  | 200    | 2  | B42148      |
| 28         | 486           | 42.3  | 202    | 2  | S72515      |
| 29         | 486           | 42.3  | 203    | 2  | JC4105      |

|    |       |      |     |   |        |                    |
|----|-------|------|-----|---|--------|--------------------|
| 30 | 485   | 42.2 | 196 | 2 | PS0279 | GTP-binding protei |
| 31 | 484.5 | 42.1 | 215 | 2 | S57474 | GTP-binding protei |
| 32 | 483   | 42.0 | 221 | 2 | H71444 | GTP-binding protei |
| 33 | 481.5 | 41.9 | 224 | 2 | T33855 | hypothetical prote |
| 34 | 479.5 | 41.7 | 204 | 2 | JC7589 | Sec4p homolog - ye |
| 35 | 479.5 | 41.7 | 208 | 2 | A38202 | GTP-binding protei |
| 36 | 478.5 | 41.6 | 206 | 2 | T14391 | GTP-binding protei |
| 37 | 478   | 41.6 | 201 | 2 | S06147 | GTP-binding protei |
| 38 | 478   | 41.6 | 203 | 2 | JC1247 | GTP-binding protei |
| 39 | 477.5 | 41.5 | 205 | 2 | S38339 | GTP-binding protei |
| 40 | 476.5 | 41.4 | 205 | 2 | T33781 | hypothetical prote |
| 41 | 476   | 41.4 | 217 | 2 | S36365 | GTP-binding protei |
| 42 | 473.5 | 41.2 | 203 | 2 | A49647 | GTP-binding protei |
| 43 | 468.5 | 40.7 | 201 | 2 | S39565 | GTP-binding protei |
| 44 | 463.5 | 40.3 | 203 | 2 | S51495 | GTP-binding protei |
| 45 | 459.5 | 40.0 | 203 | 2 | S30096 | GTP-binding protei |

## ALIGNMENTS

### RESULT 1

JC2528  
GTP-binding protein Rab26 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 02-Feb-2001  
C:Accession: JC2528  
R:Wagner, A.C.C.; Strowski, M.Z.; Goetze, B.; Williams, J.A.  
Biochem. Biophys. Res. Commun. 207, 950-956, 1995  
A:Title: Molecular cloning of a new member of the Rab protein family, Rab26, from rat  
A:Reference number: JC2528; MUID:95169156; PMID:7864900  
A:Accession: JC2528  
A:Molecule type: mRNA  
A:Residues: 1-190<WAG>  
A:Cross-references: GB:T08521  
A:Note: The authors translated the codon GTG for residue 49 as Leu, GAA for residue 1  
C:Genetics:  
A:Gene: Rab26  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:4-11/Region: nucleotide-binding motif A (P-loop)  
F:52-58/Region: GTP binding #status predicted  
F:111-114/Region: GTP binding #status predicted  
F:139-143/Region: GTP binding #status predicted

|                       |                  |  |           |             |
|-----------------------|------------------|--|-----------|-------------|
| Query Match           | 64.5%            | Score 742;   | DB 2;     | Length 190; |
| Best Local Similarity | 72.9%;           | Pred. No. 3.4e-57;   |           |             |
| Matches 137;          | Conservative 28; | Mismatches 23;   | Indels 0; | Gaps 0;     |
| QY                    | 33               | MLIGDTGVGKTCFLIQKDGAFISGTFATVGVDFRNKVVTVGVVVKVQLQIWDTAGQERF  | 92        |             |
| Db                    | 1                | MLVGDGSGVGTCLLVRFKDGAFISGTFATVGVDFRNKVVTVGVVVKVQLQIWDTAGQERF | 60        |             |
| QY                    | 93               | RSVTHAYRDAQAALLLYDITNKSSFDNIRAWLFEIHEYAQRDVVIMLLGNKADMSERV   | 152       |             |
| Db                    | 61               | RSVTHAYRDAHAALLLYDITNKSSFDNIQAWLFEIHEYAQQDVVLLMLGNKVDSTQERV  | 120       |             |
| QY                    | 153              | IRSDGETLAREYGVPELETSAKTCGMNVFLAFLAKELKYRAGHQADEPSFQIRDYVES   | 212       |             |
| Db                    | 121              | VKRDGEKLAKEYGLPFMETSAKSGLVNVDLAFLAKELKQSTKAPSEPRFLHDYVKR     | 180       |             |
| QY                    | 213              | QKRKSSCC   | 220       |             |
| Db                    | 181              | EGRGVSCC   | 188       |             |

### RESULT 2

Ti5123  
hypothetical protein W01H2.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
C:Accession: Ti5123  
R:Minx, P.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of C. elegans cosmid W01H2.  
A:Reference number: Z18296

A:Accession: T15123

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-196

A:Cross-references: EMBL:AF000192; MID:g1946982; PID:g1946985; PIDN:AAB52888.1; GSPDB:GN

A:Experimental source: strain Bristol N2; clone W01H2

C:Genetics:

A:Gene: CESP:W01H2.3

A:Map position: X

A:Introns: 41/3; 55/3; 95/3; 134/3; 160/2

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:3-119/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 533; DB 2; Length 196;

Best Local Similarity 54.5%; Pred. No. 5.2e-39;

Matches 109; Conservative 34; Mismatches 41; Indels 16; Gaps 5;

QY 31 KYMLGDTGKVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQIWDTAGOE 90

DB 4 KYMLGDSCTGTCULLIRYKOGAFNNFISTVGIDYRNKLLITMGDKKVKLIQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 150

DB 64 RFRSVTTSYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 121

QY 151 RVIREDEGTAREYGVPPLETSAKTMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

DB 122 RAVPTDEGKRLAEAYQIPEFMETSAKTGFNVDRFLGLAERMLKLYGVFPGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

DB 181 -----VADTKKPEIARCTF 195

RESULT 3

QY 31 KYMLGDTGKVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQIWDTAGOE 90

DB 4 KYMLGDSCTGTCULLIRYKOGAFNNFISTVGIDYRNKLLITMGDKKVKLIQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 150

DB 64 RFRSVTTSYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 121

QY 151 RVIREDEGTAREYGVPPLETSAKTMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

DB 122 RAVPTDEGKRLAEAYQIPEFMETSAKTGFNVDRFLGLAERMLKLYGVFPGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

DB 181 -----VADTKKPEIARCTF 195

RESULT 3

QY 31 KYMLGDTGKVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQIWDTAGOE 90

DB 4 KYMLGDSCTGTCULLIRYKOGAFNNFISTVGIDYRNKLLITMGDKKVKLIQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 150

DB 64 RFRSVTTSYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 121

QY 151 RVIREDEGTAREYGVPPLETSAKTMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

DB 122 RAVPTDEGKRLAEAYQIPEFMETSAKTGFNVDRFLGLAERMLKLYGVFPGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

DB 181 -----VADTKKPEIARCTF 195

RESULT 3

QY 31 KYMLGDTGKVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQIWDTAGOE 90

DB 4 KYMLGDSCTGTCULLIRYKOGAFNNFISTVGIDYRNKLLITMGDKKVKLIQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 150

DB 64 RFRSVTTSYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 121

QY 151 RVIREDEGTAREYGVPPLETSAKTMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

DB 122 RAVPTDEGKRLAEAYQIPEFMETSAKTGFNVDRFLGLAERMLKLYGVFPGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

DB 181 -----VADTKKPEIARCTF 195

RESULT 3

QY 31 KYMLGDTGKVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQIWDTAGOE 90

DB 4 KYMLGDSCTGTCULLIRYKOGAFNNFISTVGIDYRNKLLITMGDKKVKLIQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 150

DB 64 RFRSVTTSYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 121

QY 151 RVIREDEGTAREYGVPPLETSAKTMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

DB 122 RAVPTDEGKRLAEAYQIPEFMETSAKTGFNVDRFLGLAERMLKLYGVFPGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

QY 15 APERSPPSPSYDLTGKVMMLGDTGVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTV 74  
DB 3 SPATNKPA--AYDFLVKLLIGDSGVKSCLLRFSDGSF--TPSFATATIGIDFKIRTIEL 59  
QY 75 DGVVRVQLQIWDTAGOERFRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQ 134  
DB 60 EGKRIKLIQIWDTAGOERFTITAYYRGAMGILLVDYDDEKSFSGIRNIRNIEHQASH 119  
QY 135 DVVIMLLGNKADMSRVSIRSEGETLAREYGVPPLETSAKTMNVELAFIAIAKELKYR 194  
DB 120 SVNKMILGNKCDMTEKKVVDSSRGRSLADEYGIKELETSAKNSVNVVEAFIGLAKDIKR 179  
QY 195 AGHQADEPSFQIRDYVESQKRSSCC 220  
DB 180 MIDTPNDPDHTI--CITPNRNKNTCC 203

RESULT 4

I78851

GTP-binding protein MEL - mouse

N:Alternate names: gene MEL protein

C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 02-Feb-2001

C:Accession: I78851

R:Nimmo, E.R.; Sanders, P.G.; Padua, R.A.; Hughes, D.; Williamson, R.; Johnson, K.J.

Oncogene 6, 1347-1351, 1991

A:Title: The MEL gene: a new member of the RAB/YPT class of RAS-related genes.

A:Reference number: 158355; MUID:91360267; PMID:1886711

A:Accession: I78851

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-206 <RES>

A:Cross-references: GB:S53270; MID:g234747; PIDN:AAB19682.1; PID:g234748

C:Genetics:

A:Gene: MEL

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.9%; Score 516; DB 2; Length 206;

Best Local Similarity 49.5%; Pred. No. 1.7e-37;

Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;

QY 25 SYDLTGKVMMLGDTGVGKTCFLIOKDGAFISCTFIATVGDIFRNKVVTVGVVRKLIQI 84

DB 4 TYDYLKLLIGDSGVGKTCVLFPSSEDAF--NSTFISIGIDFKIRTIELDKRILQLIW 62

QY 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSDFNIRAWLTHEIYRDAQROVVMILLGNK 144

DB 63 DTAGQERFTITAYYRGAMGIMLVYDITNKSDFNIRNIRNIEHSADEKVMILGNK 122

QY 145 ADMSSRVSIRSEGETLAREYGVPPLETSAKTMNVELAFIAIAKELKYRAGHQ----AD 200

DB 123 CDVNDKQVSKERGEKTLADYGIKFMETSAKININVENAFETLARDIKAKMDKNWKATAA 182

QY 201 EPSFOIRDYVESQKRSSCC--CSFM 223

DB 183 GSSHGKVTVEQQRKTSFFRCSLL 206

RESULT 5

B49647

GTP-binding protein rab8 - human

C:Species: Homo sapiens (man)

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Jan-2001

C:Accession: B49647; S36817

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Helliou, R.; Tavittian, A.; Lou

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c



A:Reference number: A49647; MUID:94124602; PMID:8294494

A:Accession: B49647

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-207 <ZAH>

A:Cross-references: EMBL:X56741; NID:9452317; PIDN:CAA40065.1; PID:9452318

R:Joberty, G.; Tavittian, A.; Zahraoui, A.

FEBS Lett. 330, 323-328, 1993

A:Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.

A:Reference number: S36817; MUID:93387463; PMID:8375503

A:Accession: S36817

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 175-186 <JOB>

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (p-loop)

F:121-124/Region: nucleotide-binding motif A (p-loop)

F:151-153/Region: GTP-binding NKXD motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQFDGAFSLGTFATVGDIDFNKVVVDGVRVKLIQW 84

Db 4 TYDLFKLLIGDSGVGKTCVLFPSSEDAF-NSTFTSTIGIDFKIRTFIELDGKRKLIQW 62

Qy 85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKKSFNDINRAMLTETHEYAQRDVMILGNK 144

Db 63 DTAGQERFRITTAIYRGAMGIMLVYDITNEKSFNDINRWIRNIEEHASADVEKMLGNK 122

Qy 145 ADMSSERVISEDETLAREYGVFPFLETSKTNKNNVELAFIAIAKELKYRAGHQA--- 200

Db 123 CDVNDKROVSKERGEKALDYGIKFMETSAKINVENAFFTLARDIKAKMDKRLGNP 182

Qy 201 EPSFQIRDYVESQKRSS 218

Db 183 QGSNGVKITPDQQRSS 200

RESULT 6

B36364

GTP-binding protein rab8 - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001

R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerlial, M.

Mol. Cell. Biol. 10, 6578-6585, 1990

A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.

A:Reference number: A36364; MUID:91061765; PMID:2123294

A:Accession: B36364

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-207 <CHA>

A:Cross-references: GB:X56385; NID:920; PIDN:CAB56776.1; PID:96006436

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (p-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQFDGAFSLGTFATVGDIDFNKVVVDGVRVKLIQW 84

Db 4 TYDLFKLLIGDSGVGKTCVLFPSSEDAF-NSTFTSTIGIDFKIRTFIELDGKRKLIQW 62

Qy 85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKKSFNDINRAMLTETHEYAQRDVMILGNK 144

Db 63 DTAGQERFRITTAIYRGAMGIMLVYDITNEKSFNDINRWIRNIEEHASADVEKMLGNK 122

Qy 145 ADMSSERVISEDETLAREYGVFPFLETSKTNKNNVELAFIAIAKELKYRAGHQA--- 200

Db 123 CDVNDKROVSKERGEKALDYGIKFMETSAKINVENAFFTLARDIKAKMDKRLGNP 182

Qy 201 EPSFQIRDYVESQKRSS 218

Db 183 QGSNGVKITPDQQRSS 200

RESULT 7

T14565

GTP-binding protein - beet

N:Alternate names: small G protein

C:Species: Beta vulgaris (beet)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000

C:Accession: T14565

R:Dallery, E.; Quilef, S.; Ben Jilany, K.E.; Kerckaert, J.; Hagege, D.

A:Description: Molecular cloning and structural analysis of cDNAs that encode three s

A:Reference number: 218142

A:Accession: T14565

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-215 <DAL>

A:Cross-references: EMBL:Z49152; NID:974775; PID:974776

A:Experimental source: strain D100 KS 38080

C:Function:

A:Description: GTP-binding

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFDGAFSLGTFATVGDIDFNKVVVDG 76

Db 3 APPARADYDLIKLLIGDSGVGKSCLLRFSGSFTT-SFTTTTGIDFKIRTFIELDG 61

Qy 77 VRVKLIQWDTAGQERFRSVTHAYYRDAQALLLLYDITNKKSFNDINRAMLTETHEYAQRD 136

Db 62 KRIQLQWDTAGQERFRITTAIYRGAMGIMLVYDITNEKSFNDINRWIRNIEEHASDNV 121

Qy 137 VIMELGNKADM-SSERVISEDETLAREYGVFPFLETSKTNKNNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVVEVFFSIRARDIKORL 181

Qy 196 G---HQADEPFSFOIRDYVESQKR---RSSCC 220

Db 182 ADSITROFAQPSITIKPADQSGNQAAKSACC 213

RESULT 8

T48378

GTP-binding protein-like - Arabidopsis thaliana

N:Alternate names: protein F12E4.300

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000

C:Accession: T48378

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke,

submitted to the Protein Sequence Database, March 2000

A:Reference number: 224492

A:Accession: T48378

A:Status: preliminary

A:Molecule type: cDNA

A:Residues: 1-216 <BEV>

A:Cross-references: EMBL:AL162751

A:Experimental source: cultivar Columbia; BAC clone F12E4

Db 231 QPAGNNARPPTVQIRGQPVAK--NGCCS 257

Db 231 QPAGNNARPPTVQIRGQVPAQK--NGCCS 257

QY 26 YDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIIDFRNKVVTVDGVRVKLQIWD 85

QY 26 YDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIIDFRNKVVTVDGVRVKLQIWD 85

```

Db      6 YDMLFKLLIGDSGVKTCILYFSDAP-NTTIFSTIGIDFKIKIYELGKKIKQIWD 64
QY      86 TAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEHYAQRDVVIMLGNKA 145
Db      65 TAGOERFHTTTSYRGAMGIMLVYDITNKSFDNIAKWLNRIDHASEDVVKILGNKC 124
QY      146 DMSERSVIRSEDETALAREYGVFPLETSAGTKGMNVELAFIAIAKELKYRAGHODEPSFQ 205
Db      125 DMSDRVSVRERKEAQDHGIFSHETSAKLNHVDTAFYDLAEAILAKMPDSTDQSRD 184
QY      206 IRDYVESQKRSS--CC 220
Db      185 TVNPVQPRQSSGGCC 201

RESULT 12
S38740
GTP-binding protein - rice
N:Alternate names: ras-related small GTP-binding protein
C:Species: Oryza sativa (rice)
C>Date: 27-May-1994 #sequence_revision 27-Feb-1997 #text_change 02-Feb-2001
C:Accession: S38740
R:Kidou, S.; Anal, T.; Umeda, M.; Aotsuka, S.; Tsuge, T.; Kato, A.; Uchimiya, H.
FEBS Lett. 332, 282-286, 1993
A:Title: Molecular structure of ras-related small GTP-binding protein genes of rice plant
A:Reference number: S38740; MUID:94009718; PMID:8405471
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <RID>
A:Cross-references: GB:S66160; NID:9432606; PIDN:AB28535.1; PID:9432607
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 43.5%; Score 500; DB 2; Length 202;
Best Local Similarity 46.3%; Pred. No. 4e-36;
Matches 94; Conservative 49; Mismatches 54; Indels 6; Gaps 3;

QY      23 SPYDITGKVMLLGDTGVGKTCFLIQKDGAFISGTFIATVGIDFRNKVVTVGVRVKLQ 82
Db      2 NPEYDYLFLKLLIGDSGVKSCILLRFADDSYLE-SYISTIGVDKIRTVEQDGKTIKLQ 60
QY      83 IWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEHYAQRDVVIMLGL 142
Db      61 IWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEHYAQRDVVIMLGL 120
QY      143 NKADMSERSVIRSEDETALAREYGVFPLETSAGTKGMNVELAFIAIAKELKYRAGHQ---- 198
Db      121 NKCDLAENRVSVYEGAKALADEIGIPFLETSAKDINVEKAFMTMAGEIKNRASQGRTN 180
QY      199 ADPEPSQIRDYVESQKRSSCCS 221
Db      181 ASKPA-TVQMPRPVQAQSSCCS 202

RESULT 13
T45901
GTPase AtRAB8 - Arabidopsis thaliana
N:Alternate names: protein F4P12.310
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T45901
R:Blöcker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223016
A:Accession: T45901
A:Status: preliminary

```

```

A:Molecule type: DNA
A:Residues: 1-216 <BLO>
A:Cross-references: EMBL:AL132966
C:Experimental source: cultivar Columbia; BAC clone F4P12
C:Genetics:
A:Map position: 3
A:Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3
A:Note: F4P12.310
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
F:16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 43.5%; Score 500; DB 2; Length 216;
Best Local Similarity 46.5%; Pred. No. 4.3e-36;
Matches 99; Conservative 45; Mismatches 57; Indels 12; Gaps 5;

QY      19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFISGTFIATVGIDFRNKVVTVDG 76
Db      3 APPARADYDYLKLLIGDSGVKSCILLRSDGSFTT-SFTTIGIDFKIRTIELDG 61
QY      77 VRVKLIWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEHYAQRDV 136
Db      62 KRILQIWDTAGOERFRFTITAYYRGAMGILLVYDVTDESSENNIRNIRNIEGHASDSV 121
QY      137 VIMLGNKADM-SSEVRSEDETALAREYGVFPLETSAGTKGMNVELAFIAIAKELKYRA 195
Db      122 NKILVGNKADMDESKRAVPKSGQALADEYGMKFFETSAKTNLNVVEEVFFSIADIKQRL 181
QY      196 GH---QADEPSFOIRDY-----VESQKRSSCC 220
Db      182 ADTDARAEPQFIKINQSDGAGTSQATQKSACC 214

RESULT 14
A34716
GTP-binding protein SAS1 - allme mold (Dictyostellium discoideum)
C:Species: Dictyostellium discoideum
C>Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 02-Feb-2001
C:Accession: A34716; A61571
R:Saxe, S.A.; Kimmel, A.R.
Mol. Cell. Biol. 10, 2367-2378, 1990
A:Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostellium discoideum with se
A:Reference number: A34716; MUID:90220623; PMID:2109188
A:Accession: A34716
A:Molecule type: mRNA
A:Residues: 1-208 <SAX>
A:Cross-references: GB:M34456
R:Saxe, S.A.; Kimmel, A.R.
Dev. Genet. 9, 255-265, 1988
A:Title: Genes encoding novel GTP-binding proteins in Dictyostellium.
A:Reference number: A61571; MUID:89209367; PMID:3149563
A:Accession: A61571
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-208 <SA2>
C:Genetics:
A:Gene: SAS1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F:16-131/Domain: translation elongation factor Tu homology <ETU>
F:22-29/Region: nucleotide-binding motif A (P-loop)
F:128-131/Region: GTP-binding NKXD motif
F:158-160/Region: GTP-binding SAK/L motif
F:207,208/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 43.3%; Score 498; DB 2; Length 208;
Best Local Similarity 46.6%; Pred. No. 6.2e-36;
Matches 95; Conservative 49; Mismatches 50; Indels 10; Gaps 3;

QY      23 SPYDITGKVMLLGDTGVGKTCFLIQKDGAFISGTFIATVGIDFRNKVVTVGVRVKLQ 82
Db      9 SAAYDYLIKLLIGDSGVKSCILLRSEDSF-TPSFTTIGIDFKIRTIELGKRIKLQ 67
QY      83 IWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEHYAQRDVVIMLGL 142

```

Db 68 IWDTAGQERFTITAYTRGAMGILLVYDVYDEKSFNGIRNWIIEQHATDSYNNKMLIG 127  
QY 143 NKADSSSRVIRSEDETLAREYGVPPLETSAKTGMVNLAFALAKELKYRAGHQADEP 202  
Db 128 NKCDMAEKVVDSSRGKSLADEYGIKLETSAKNSINVEEAFISLANDIKKR---MIDTP 184  
QY 203 SFQIR-----DYVESQKRSSCC 220  
Db 185 NEQPOVQPGTNLGANNKKKACC 208

## RESULT 15

S41430  
GTP-binding protein, ras-like (clone vfa-ypt1) - fava bean  
N:Alternate names: guanine nucleotide regulatory protein  
C:Species: Vicia faba (fava bean)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 02-Feb-2001  
C:Accession: S41430  
R:Saalbach, G.; Thielmann, J.  
submitted to the EMBL Data Library, January 1994  
A:Description: Sequences of cDNA clones from cotyledons of Vicia faba encoding ypt/rab-  
A:Reference number: S41430  
A:Accession: S41430  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-202 <SAA>  
A:Cross-references: EMBL:Z29590; NID:g452358; PIDN:CAA82707.1; PID:g452359  
A:Experimental source: clone vfa-ypt1  
A:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop  
F:9-124/Domain: translation elongation factor Tu homology <ETU>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:151-153/Region: GTP-binding SAK/L motif  
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 43.3%; Score 497.5; DB 2; Length 202;  
Best Local Similarity 45.8%; Pred. NO. 6.6e-36;  
Matches 93; Conservative 46; Mismatches 57; Indels 7; Gaps 3;  
QY 23 SPSYDLTGKVMLLGDTGKTCFLIQKDGAFLSGTATVGTDFRNKVVTVDGVYVKLQ 82  
Db 2 NPEYDYLKLLIGDSGVGKSCLLRFADDSYID-SYISTIGVDFKIRTVEDQGTIRLQ 60  
QY 83 IWDTAGQERFRSVTHAYTRDAQALLLYDITNKSFNIRAWLTIHEYAQRDVYIMLLG 142  
Db 61 IWDTAGQERFTITSSYYRGAGHIIYVDVDESEFNKQWLSIDRYASDNVKNLLVG 120  
QY 143 NKADSSSRVIRSEDETLAREYGVPPLETSAKTGMVNLAFALAKELKYRAGHQ----198  
Db 121 NKCDUTENRVPYETAKAFADDEIGIPFMETSAKOSTNVEQAFMAMASSIKERMASQPTNN 180  
QY 199 ADEPSFQIRDYVESQKRSSCCS 221  
Db 181 ARPPTVQIRGQPVGOK--SGCCS 201

Search completed: June 18, 2003, 15:49:08  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 20:56:44 ; Search time 6785 Seconds  
(without alignments)  
11469.551 Million cell updates/sec

Title: US-09-817-199b-1

Perfect score: 2674

Sequence: 1 ttccgtcgggccggcact.....aaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_hgt:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_hgt\_hum:\*
- 31: em\_hgt\_inv:\*
- 32: em\_hgt\_other:\*
- 33: em\_hgt\_mus:\*
- 34: em\_hgt\_pln:\*
- 35: em\_hgt\_rod:\*
- 36: em\_hgt\_mam:\*
- 37: em\_hgt\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgtgo\_hum:\*
- 40: em\_hgtgo\_mus:\*
- 41: em\_hgtgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID         | Description        |
|------------|--------|-------------|--------|----|------------|--------------------|
| 1          | 2161.8 | 80.8        | 2195   | 9  | AK057069   | AK057069 Homo sapi |
| 2          | 1962   | 73.4        | 1977   | 9  | AK098068   | AK098068 Homo sapi |
| 3          | 1354.8 | 50.7        | 3105   | 9  | AK054846   | AK054846 Homo sapi |
| 4          | 1086   | 40.6        | 1116   | 6  | AX236082   | AX236082 Sequence  |
| 5          | 938.4  | 35.1        | 1700   | 9  | BC016615   | BC016615 Homo sapi |
| 6          | 576    | 21.5        | 690    | 6  | AX236084   | AX236084 Sequence  |
| 7          | 563.6  | 21.1        | 576    | 10 | AF233582   | AF233582 Mus muscu |
| 8          | 315.6  | 11.8        | 1513   | 9  | BC007681   | BC007681 Homo sapi |
| 9          | 312.4  | 11.7        | 1098   | 10 | RN018771   | U18771 Rattus norv |
| 10         | 304.6  | 11.4        | 1320   | 9  | AB027137   | AB027137 Homo sapi |
| 11         | 303.2  | 11.3        | 573    | 9  | AF498952   | AF498952 Homo sapi |
| 12         | 292.6  | 10.9        | 1340   | 6  | AR062279   | AR062279 Sequence  |
| 13         | 234.2  | 8.8         | 2536   | 3  | AY061826   | AY061826 Drosophil |
| 14         | 214.8  | 8.0         | 127587 | 9  | HS941F9    | 295331 Human DNA s |
| 15         | 214.2  | 8.0         | 102332 | 9  | AC010352   | AC010352 Homo sapi |
| 16         | 214.2  | 8.0         | 107465 | 9  | AC010434   | AC010434 Homo sapi |
| 17         | 212.8  | 8.0         | 146740 | 9  | CNS01DIX   | AL132989 Human chr |
| 18         | 212.6  | 8.0         | 174309 | 9  | AC016656   | AC016656 Homo sapi |
| 19         | 212.6  | 8.0         | 176640 | 9  | AC016652   | AC016652 Homo sapi |
| 20         | 211.6  | 7.9         | 148841 | 9  | AC011462   | AC011462 Homo sapi |
| 21         | 210.4  | 7.9         | 199601 | 9  | AL389915   | AL389915 Human DNA |
| 22         | 210.2  | 7.9         | 90000  | 9  | AC117467   | AC117467 Homo sapi |
| 23         | 209.6  | 7.8         | 141771 | 9  | AC091821   | AC091821 Homo sapi |
| 24         | 209    | 7.8         | 110000 | 2  | AC013391_0 | AC013391 Homo sapi |
| 25         | 209    | 7.8         | 125980 | 9  | AC093223   | AC093223 Homo sapi |
| 26         | 209    | 7.8         | 157141 | 9  | AC011470   | AC011470 Homo sapi |
| 27         | 209    | 7.8         | 176304 | 9  | AC013787   | AC013787 Homo sapi |
| 28         | 208.4  | 7.8         | 167509 | 2  | AC012107   | AC012107 Homo sapi |
| 29         | 208    | 7.8         | 43790  | 9  | AC113388   | AC113388 Homo sapi |
| 30         | 208    | 7.8         | 86685  | 9  | AC118457   | AC118457 Homo sapi |
| 31         | 208    | 7.8         | 107995 | 2  | HSJ356J11  | AL121971 Homo sapi |
| 32         | 208    | 7.8         | 172048 | 9  | HS179N16   | 295152 Homo sapien |
| 33         | 207.6  | 7.8         | 137271 | 9  | AC004552   | AC004552 Homo sapi |
| 34         | 207.6  | 7.8         | 207366 | 9  | AC090001   | AC090001 Homo sapi |
| 35         | 207.4  | 7.8         | 136222 | 9  | AC004099   | AC004099 Homo sapi |
| 36         | 207.2  | 7.7         | 63905  | 9  | AC022515   | AC022515 Homo sapi |
| 37         | 207.2  | 7.7         | 186330 | 9  | AC007376   | AC007376 Homo sapi |
| 38         | 207.2  | 7.7         | 271814 | 2  | AC027406   | AC027406 Homo sapi |
| 39         | 207    | 7.7         | 154289 | 2  | AC027714   | AC027714 Homo sapi |
| 40         | 207    | 7.7         | 161757 | 2  | AC044901   | AC044901 Homo sapi |
| 41         | 207    | 7.7         | 183528 | 2  | AC130416   | AC130416 Homo sapi |
| 42         | 207    | 7.7         | 187399 | 9  | AC090970   | AC090970 Homo sapi |
| 43         | 207    | 7.7         | 209797 | 2  | AC005748   | AC005748 Homo sapi |
| 44         | 207    | 7.7         | 346940 | 2  | AC107420   | AC107420 Homo sapi |
| 45         | 206.8  | 7.7         | 157113 | 9  | AC083797   | AC083797 Homo sapi |

ALIGNMENTS

RESULT 1  
AK057069  
LOCUS AK057069  
DEFINITION Homo sapiens cDNA FLJ32507 fis, clone SMINT1000048, moderately similar to Mus musculus GTPase Rab37 mRNA.  
ACCESSION AK057069  
VERSION AK057069.1 GI:16552648  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens small intestine cDNA to mRNA, clone\_lib:SMINT1 clone:SMINT1000048.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,

AK057069 2195 bp mRNA linear PRI 01-AUG-2002  
Homo sapiens cDNA FLJ32507 fis, clone SMINT1000048, moderately similar to Mus musculus GTPase Rab37 mRNA.

Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,  
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fuji,A.,  
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.  
and Isogai,T.  
NEDO human cDNA sequencing project  
Unpublished

# TITLE JOURNAL REFERENCE

2 (bases 1 to 2195)  
Isogai,T., Otsuki,T. and Sugiyama,T.  
Direct Submission  
Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass. sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

## FEATURES

### source

Location/Qualifiers

1..2195  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SMINT1000048"  
/tissue\_type="small intestine"  
/clone\_lib="SMINT1"  
/note="cloning vector: pME18SFL3"

BASE COUNT 511 a 596 c 582 g 506 t

Query Match 80.8%; Score 2161.8; DB 9; Length 2195;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 472 AGCGGATATGACGAGGAAAGAGTGATCCGTTCCGAGACGAGAGACCTTGGCCAGGG 531  
DB 31 AGCGGATATGACGAGGAAAGAGTGATCCGTTCCGAGACGAGAGACCTTGGCCAGGG 90  
QY 532 AGTAGCGGTTCCTTCCTGGAGACCGGCCAAGACTGGCATGAATGGAGTTAGCCT 591  
DB 91 AGTAGCGGTTCCTTCCTGGAGACCGGCCAAGACTGGCATGAATGGAGTTAGCCT 150  
QY 592 TTCTGGCCATGCCAAGGACTGAATACCGGGCCGCGCATCAGCGGATGAGCCAGCT 651  
DB 151 TTCTGGCCATGCCAAGGACTGAATACCGGGCCGCGCATCAGCGGATGAGCCAGCT 210  
QY 652 TCCAGATCCGAGACTATGTAGAGTCCAGAGAGCGCTCCAGCTGCTCTCTTCATGT 711  
DB 211 TCCAGATCCGAGACTATGTAGAGTCCAGAGAGCGCTCCAGCTGCTCTCTTCATGT 270  
QY 712 GAATCCAGGGGGCAGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCTCC 771  
DB 271 GAATCCAGGGGGCAGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCTCC 330  
QY 772 AGGCTGCTTATTCAGAGGCTGAGCCATGGGGAAAGATGGAGACTCACTGCAC 831  
DB 331 AGGCTGCTTATTCAGAGGCTGAGCCATGGGGAAAGATGGAGACTCACTGCAC 390  
QY 832 AGCGCTTCCTAGCAGGAGGTATACCTCAACTCTACTTCTAGTTCCCTGCGGTCTCCCG 891  
DB 391 AGCGCTTCCTAGCAGGAGGTATACCTCAACTCTACTTCTAGTTCCCTGCGGTCTCCCG 450  
QY 892 CATCCACAGGAGGGTAAACACTTATTTTAAATAGTACATAATTAATACCAA 951  
DB 451 CATCCACAGGAGGGTAAACACTTATTTTAAATAGTACATAATTAATACCAA 510  
QY 952 AAAAGGCGCTTGGATCCCAAAAACGAGGCTGGAGCTAGTGGCCCTTTTGTCTTCTA 1011  
DB 511 AAAAGGCGCTTGGATCCCAAAAACGAGGCTGGAGCTAGTGGCCCTTTTGTCTTCTA 570

QY 1012 GGACTTGGGGGGCGGCGCTCCTCCTTAAGCATAAAGAGTGGTGTCTCCAGCTCAG 1071  
DB 571 GGACTTGGGGGGCGGCGCTCCTCCTTAAGCATAAAGAGTGGTGTCTCCAGCTCAG 630  
QY 1072 CCCAGGGGACACAGATGACATTTGGGGTGGAGGAGTAAATGACTCATGCGACCCCTC 1131  
DB 631 CCCAGGGGACACAGATGACATTTGGGGTGGAGGAGTAAATGACTCATGCGACCCCTC 690  
QY 1132 AGTTCAAGCTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCTCTCCCTCTCCAGGAG 1191  
DB 691 AGTTCAAGCTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCTCTCCCTCTCCAGGAG 750  
QY 1192 CTTATCTTCGCCCCATCTCCCAATAAAGTGGGCCCTTGTGTGTAGGAGAACCAAGCC 1251  
DB 751 CTTATCTTCGCCCCATCTCCCAATAAAGTGGGCCCTTGTGTGTAGGAGAACCAAGCC 810  
QY 1252 TCAGGGAAGATGAAGATATGAGATGGAGGGGGAGGAGGAGGAGGAGAGTAGGTGTC 1311  
DB 811 TCAGGGAAGATGAAGATATGAGATGGAGGGGGAGGAGGAGGAGGAGAGTAGGTGTC 870  
QY 1312 TAGCTGGCTATCTCTGGCCTTACTAACACCCCTCGAGGCAATGCCCTTTTCTCCAGCA 1371  
DB 871 TAGCTGGCTATCTCTGGCCTTACTAACACCCCTCGAGGCAATGCCCTTTTCTCCAGCA 930  
QY 1372 CACAAGCAGATTTGGGGACCTGGAAATATTGGTTCAGGCTCCTGTCTCTGGACTTCAG 1431  
DB 931 CACAAGCAGATTTGGGGACCTGGAAATATTGGTTCAGGCTCCTGTCTCTGGACTTCAG 990  
QY 1432 ATCTCTGGGGAGCCCTCCCGCCCTGAATCCTCGCTTAGCTACCTCTCCCTGTGCA 1491  
DB 991 ATCTCTGGGGAGCCCTCCCGCCCTGAATCCTCGCTTAGCTACCTCTCCCTGTGCA 1050  
QY 1492 CCTAAAAACCTCAGTCAAGACTAGGAAAGAGTTTGTATTTTATTTTGAATGGAG 1551  
DB 1051 CCTAAAAACCTCAGTCAAGACTAGGAAAGAGTTTGTATTTTATTTTGAATGGAG 1110  
QY 1552 TCTCTCTCTCTCGCCAGGCTCAGTGTAGTGTGCAATCTCCGCTCACTCAACACTCA 1611  
DB 1111 TCTCTCTCTCTCGCCAGGCTGAGTGTAGTGTGCAATCTCCGCTCACTCAACACTCA 1170  
QY 1612 CTCCCTGGGGCTCAAGCGATCTCCAGCTCAGCGCCGAGTAGCTGGAGCTATAGTG 1671  
DB 1171 CTCCCTGGGGCTCAAGCGATCTCCAGCTCAGCGCCGAGTAGCTGGAGCTATAGTG 1230  
QY 1672 TGTACCATCAGCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCCGCTATGTTG 1731  
DB 1231 TGTACCATCAGCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCCGCTATGTTG 1290  
QY 1732 CCCAGGCTGGTCTGAATTCCTGAGCTCAAGCAACTCCGCGCTCGGCTCCCAAGTA 1791  
DB 1291 CCCAGGCTGGTCTGAATTCCTGAGCTCAAGCAACTCCGCGCTCGGCTCCCAAGTA 1350  
QY 1792 CTGGGATTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1851  
DB 1351 CTGGGATTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1410  
QY 1852 GCAGGATGAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1911  
DB 1411 GCAGGATGAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1470  
QY 1912 TTAAGAGCCAGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1971  
DB 1471 TTAAGAGCCAGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1530  
QY 1972 AGAATCCCTGGCTCCTGGAGCTGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2031  
DB 1531 AGAATCCCTGGCTCCTGGAGCTGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1590  
QY 2032 TATCCCTTTGAGCCAGCTGCTGCTCCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2091  
DB 1591 TATCCCTTTGAGCCAGCTGCTGCTCCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 1650  
QY 2092 CTCAAGCAGATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2151

```

Db 1651 CTCACACAGATACACAGGCGCTAAACTGCTTTACTCTCCCTCCTACTGAGTCAGGTAGGT 1710
Qy 2152 GGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGATAGTAACAAAGTCATGT 2211
Db 1711 GGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGATAGTAACAAAGTCATGT 1770
Qy 2212 GGTATGTCGTGGGTAGAGAGAGGGTAGCAAGTTCATGTGCTCTCTGTCACATATC 2271
Db 1771 GGTATGTCGTGGGTAGAGAGAGGGTAGCAAGTTCATGTGCTCTCTGTCACATATC 1830
Qy 2272 TCCCAAGAGCTGATCCTGTCATGGAAGTGGACAGGAACATGAGTCAATGACCTGCA 2331
Db 1831 TCCCAAGAGCTGATCCTGTCATGGAAGTGGACAGGAACATGAGTCAATGACCTGCA 1890
Qy 2332 GGCATCTTTACTGAGGCTCTCCGCGCTGGAGGGGAGAGGGGAGGAAGATATGCGC 2391
Db 1891 GGCATCTTTACTGAGGCTCTCCGCGCTGGAGGGGAGAGGGGAGGAAGATATGCGC 1950
Qy 2392 TGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTCTGTCGAGCATCA 2451
Db 1951 TGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTCTGTCGAGCATCA 2010
Qy 2452 CGGCTCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGAGCAGGG 2511
Db 2011 CGGCTCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGAGCAGGG 2070
Qy 2512 AGGAGGTGACATTTGCCAGTCTCTCTGCGGCCCAAGGCAGGTTGCGAGGAGATCCATCC 2571
Db 2071 AGGAGGTGACATTTGCCAGTCTCTCTGCGGCCCAAGGCAGGTTGCGAGGAGATCCATCC 2130
Qy 2572 CATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGATTAACATGTCAGTATTTGG 2631
Db 2131 CATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGATTAACATGTCAGTATTTGG 2190
Qy 2632 AAAGC 2636
Db 2191 AAAGC 2195

```

```

RESULT 2
LOCUS AK098068
DEFINITION Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
to Mus musculus GTFase Rab37 (Rab37) mRNA.
ACCESSION AK098068
VERSION AK098068.1 Gi:21758001
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens trachea cDNA to mRNA, clone_lib:TRACH2
clone:TRACH2000540.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A.,
Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K.,
Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1977)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kasuga-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:

```

```

Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source            1..1977
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TRACH2000540"
                     /tissue_type="trachea"
                     /clone_lib="TRACH2"
                     /note="cloning vector: pME18SFL3"
     CDS                57..728
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC05227.1"
                     /db_xref="GI:21758002"
                     /translation="MTGTPGAVATRDGEAPSPSPSYDLTKVMLLGGTVGKTK
FLIOFKDGAFLSGTFIATGIDFRNKVVTVGVKVLQIMDAGQERFSRVTHAYRD
AQALLLYDITKSFDFNIRAWLTIHEYAQRDVVIMLLGNKADMSRSRVSJEDGET
LAREYGVFLETSAKTGMNVELAFLAIKELKYRAGHQADPEPSQIRDYVESQKRKS
CCSPM"
BASE COUNT  442 a 560 c 539 g 436 t
ORIGIN
Query Match      73.4%; Score 1962; DB 9; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGCGCTCGGGCGGCGGCACTGCTACCTCTCGTCCAGGGACATGACGGGACGCCAGCGG 60
Db 16 TTGCGCTCGGGCGGCGGCACTGCTACCTCTCGTCCAGGGACATGACGGGACGCCAGCGG 75
Qy 61 CCGTTGCCACCGGGATGGCGAGGCGCCCGGAGCGCTCCCGCCCTGCGAGCTCGGAGCTACG 120
Db 76 CCGTTGCCACCGGGATGGCGAGGCGCCCGGAGCGCTCCCGCCCTGCGAGCTCGGAGCTACG 135
Qy 121 ACCTACGGGCAAGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTGA 180
Db 136 ACCTACGGGCAAGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTGA 195
Qy 181 TCCAAATCAAGAGCGGGCGCTTCTGTCGGAGACCTTCATAGCCACCGTCGGCATAGACT 240
Db 196 TCCAAATCAAGAGCGGGCGCTTCTGTCGGAGACCTTCATAGCCACCGTCGGCATAGACT 255
Qy 241 TCAGGAACAAGTGTGCTGCTGTTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 300
Db 256 TCAGGAACAAGTGTGCTGCTGTTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 315
Qy 301 CTGGGACGAACGGTTCCGAAGCGGTCACCCCATGCTTATTACAGAGATGCTCAGGCGCTTGC 360
Db 316 CTGGGACGAACGGTTCCGAAGCGGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTGC 375
Qy 361 TTCGCTCTATGACATCAACCAACAAATCTTCTTCGCAACATCAGGCGCTGGCTCACTG 420
Db 376 TTCGCTCTATGACATCAACCAACAAATCTTCTTCGCAACATCAGGCGCTGGCTCACTG 435
Qy 421 AGATTCAATGATGATGCTCCAGAGGAGCTGTTGATCATCTGCTAGGCAACAGGCGGATA 480
Db 436 AGATTCAATGATGATGCTCCAGAGGAGCTGTTGATCATCTGCTAGGCAACAGGCGGATA 495
Qy 481 TGAGCAGCGAAGAGTGTATCGGTTCCGAAGACGGAGAGACCTTGGCCAGGAGTACGGGTG 540
Db 496 TGAGCAGCGAAGAGTGTATCGGTTCCGAAGACGGAGAGACCTTGGCCAGGAGTACGGGTG 555
Qy 541 TTCGCTCTCTGAGACACCGGCAAGACTGGCATGAATGTGGAGTTAGCTTTCTGGCCA 600
Db 556 TTCGCTCTCTGAGACACCGGCAAGACTGGCATGAATGTGGAGTTAGCTTTCTGGCCA 615
Qy 601 TCGCCAAAGGAACTGAATACCGGCGCGGCGCATCAGGCGGATGAGCCAGCTTCCAGATCC 660

```





QY 649 GCTTCCAGATCCGAGACTATGTAGAGTCCAGAGAAGCGCTCCAGGTGCTGCTCTTCA 708  
DB 1792 GCTTCCAGATCCGAGACTATGTAGAGTCCAGAGAAGCGCTCCAGGTGCTGCTCTTCA 1851  
QY 709 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATCAGCCCTTCCGCT 768  
DB 1852 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATCAGCCCTTCCGCT 1911  
QY 769 CCAGGCTGCTTATTCAGAGGCTGAGCCATGGGAGAAAGATGGAGGACTACTG 828  
DB 1912 CCAGGCTGCTTATTCAGAGGCTGAGCCATGGGAGAAAGATGGAGGACTACTG 1971  
QY 829 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTTACTGAGTTCTGCGGTCTCC 888  
DB 1972 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTTACTGAGTTCTGCGGTCTCC 2031  
QY 889 CCGCATCCAGAGGAGGTAACACACTTACTTTTAAATAGTACATAATTAATAC 948  
DB 2032 CCGCATCCAGAGGAGGTAACACACTTACTTTTAAATAGTACATAATTAATAC 2091  
QY 949 CAAAAAGGCGCTGGATCCCCAAAAACCGAGCTGGAGCTAGTGGCCCTTTTGGCTT 1008  
DB 2092 CAAAAAGGCGCTGGATCCCCAAAAACCGAGCTGGAGCTAGTGGCCCTTTTGGCTT 2151  
QY 1009 CTAGGACTTGGGGGCGGCGCTCCCTTAAGCATAAAGGTGCTGCTCCAGCT 1058  
DB 2152 CTAGGACTTGGGGGCGGCGCTCCCTTAAGCATAAAGGTGCTGCTCCAGCT 2211  
QY 1069 CAGCCCGAGGACACAGATGCACTTTGGGGTGAGGCGAGTAATCACTCCATCGCAC 1128  
DB 2212 CAGCCCGAGGACACAGATGCACTTTGGGGTGAGGCGAGTAATCACTCCATCGCAC 2271  
QY 1129 CTCAGTTAGCTGACAGAGGCTCAGTGAGCCCGAGCTTCACTGTCTCCGCTCTCCAG 1188  
DB 2272 CTCAGTTAGCTGACAGAGGCTCAGTGAGCCCGAGCTTCACTGTCTCCGCTCTCCAG 2331  
QY 1189 GAGTTATCTGCGCCCATCTCCAAATAGTGGGCGCTTGTGCTGTGAGGAACACAA 1248  
DB 2332 GAGTTATCTGCGCCCATCTCCAAATAGTGGGCGCTTGTGCTGTGAGGAACACAA 2391  
QY 1249 GCCTCAGGGAAGATAAGAGATATGAGATGGAGGGGAGGACAGAGGAGAGTAGG 1308  
DB 2392 GCCTCAGGGAAGATAAGAGATATGAGATGGAGGGGAGGACAGAGGAGAGTAGG 2451  
QY 1309 GTCTAGCTGCTATCTCTGGCTTACTAACACCCCGCTGGAGGATGCCCTTTCTCCA 1368  
DB 2452 GTCTAGCTGCTATCTCTGGCTTACTAACACCCCGCTGGAGGATGCCCTTTCTCCA 2511  
QY 1369 GCACAAAGCACATTTGGGCGACCTGGAAATATTTGTTCCAGGCTTCTCTGGACTT 1428  
DB 2512 GCACAAAGCACATTTGGGCGACCTGGAAATATTTGTTCCAGGCTTCTCTGGACTT 2571  
QY 1429 CAGATCTTGGGGAGCCCTCCCGCCCTGAATCCCTGGCTTACTACTCTCTGCTGT 1488  
DB 2572 CAGATCTTGGGGAGCCCTCCCGCCCTGAATCCCTGGCTTACTACTCTCTGCTGT 2631  
QY 1489 GCACCTAAAACCTCAGGTGAGAACTAGGAAAAGAGTTTGTATTTTATTTTGAATG 1548  
DB 2632 GCACCTAAAACCTCAGGTGAGAACTAGGAAAAGAGTTTGTATTTTATTTTGAATG 2691  
QY 1549 GAGTCTGTTCTGTCGCGCCAGGCTGAGGTGAGTAGTGCATCTCCGCTCACTACACCT 1608  
DB 2692 GAGTCTGTTCTGTCGCGCCAGGCTGAGGTGAGTAGTGCATCTCCGCTCACTACACCT 2751  
QY 1609 CCAGTCTGCTGGGCTCAAGGATCTCCAGCTCAGCGCGGAGTAGTGGGACTATAG 1668  
DB 2752 CCAGTCTGCTGGGCTCAAGGATCTCCAGCTCAGCGCGGAGTAGTGGGACTATAG 2811  
QY 1669 GTGTGTACCATCACACTGGCTAATTTTGTATTTTGTAGACACAGGTTTCGCCATG 1728  
DB 2812 GTGTGTACCATCACACTGGCTAATTTTGTATTTTGTAGACACAGGTTTCGCCATG 2871

QY 1729 TTGCGCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCCGCTCCAAA 1788  
DB 2872 TTGCGCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCCGCTCCAAA 2931  
QY 1789 GTACTGGGATTACACGAGAGGACCATGCGCAGGCTAGATGTCTTATCCCAATCCT 1848  
DB 2932 GTACTGGGATTACACGAGAGGACCATGCGCAGGCTAGATGTCTTATCCCAATCCT 2991  
QY 1849 TTGCGCAGGCTGAGCTCCACAGGCTATTTCTTCAAGCAGCTGAAGTGTAGCCCTCT 1908  
DB 2992 TTGCGCAGGCTGAGCTCCACAGGCTATTTCTTCAAGCAGCTGAAGTGTAGCCCTCT 3051  
QY 1909 GGGTTAAGAGCCAGATAAGAGAAATCCCTTCTAGTGTGGATGTGTGTG 1962  
DB 3052 GGGTTAAGAGCCAGATAAGAGAAATCCCTTCTAGTGTGGATGTGTGTG 3105

RESULT 4  
AX236082 LOCUS AX236082 1116 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 13 from Patent WO0164887.  
ACCESSION AX236082  
VERSION AX236082.1 GI:15795889  
KEYWORDS  
SOURCE human;  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1116)  
AUTHORS Meyers, R. A.  
TITLE 23224, 27423, 32700, 32712, novel human g-proteins  
JOURNAL Patent: WO 0164887-A 13 07-SEP-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source 1..1116  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
CDS  
124..699  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC88255.1"  
/db\_xref="GI:15795890"  
/translation="MLLGDVGVKTCFLIOFKDGAFLSGFIATVIGIDFRNKVTVVG  
VRVRLQIWDTAGQERFSRTHAYTRDAQLLLDITNKSFDIRLAWLTFIHEYAQR  
DVVIMLGNKADMSSESRVISEDGETLAREYGVPLETSARTKMNVELAFIALEKL  
YRAGHQADEPSFQIRDYVESQKRSSCCSEF"  
BASE COUNT 250 a 313 c 322 g 221 t  
ORIGIN

Query Match 40.6%; Score 1086; DB 6; Length 1116;  
Best Local Similarity 99.5%; Pred. No. 5.1e-280;  
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGGACATACGGGCGACCGCAGGCGCGTTGGCCACCCCGGATGGCAGGCCCGGAGGCT 96  
DB 23 GCGTCCGAGCGGCGACCGCAGCGCGCGTTGGCCACCCCGGATGGCAGGCCCGGAGGCT 82  
QY 97 CCCGCCCTCGAGTCCGAGCTACGACCTCAGCGCAAGGTGATGCTTGTGGGAGACACAG 156  
DB 83 CCCGCCCTCGAGTCCGAGCTACGACCTCAGCGCAAGGTGATGCTTGTGGGAGACACAG 142  
QY 157 GCGTCCGCAAAACATGTTTCTGTATCCAAATTCAAAGACGGGCGCTTCTGTCCGAACCT 216  
DB 143 GCGTCCGCAAAACATGTTTCTGTATCCAAATTCAAAGACGGGCGCTTCTGTCCGAACCT 202  
QY 217 TCATAGCACCGCTCGGATAGACTTCAGGAACAAGTGGTACGTGGATGGCTGGAGAG 276  
DB 203 TCATAGCACCGCTCGGATAGACTTCAGGAACAAGTGGTACGTGGATGGCTGGAGAG 262  
QY 277 TGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAACGTTCCGAGGAGTCAACCATGCTT 336  
DB 263 TGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAACGTTCCGAGGAGTCAACCATGCTT 322

337 ATTACAGAGATGCTCAGGCTTGTCTGCTGTATGACATCACCACAAATCTCTTTG 396  
 323 ATTACAGAGATGCTCAGGCTTGTCTGCTGTATGACATCACCACAAATCTCTTTG 382  
 397 ACAACATCAGGCGCTGGCTCACTGAGATTCATGAGTATGCCAGAGGAGCGTGGTATCA 456  
 383 ACAACATCAGGCGCTGGCTCACTGAGATTCATGAGTATGCCAGAGGAGCGTGGTATCA 442  
 457 TCGTCTAGGCAACAGGCGGATNTAGCAGGAGGAAGAGTATCGCTCCGAGACGGAG 516  
 443 TCGTCTAGGCAACAGGCGGATNTAGCAGGAGGAAGAGTATCGCTCCGAGACGGAG 502  
 517 AGACCTTGGCCAGGAGTACGGTGTCTCTTCTGGAGACCGCCAGAGCTGGCATCA 576  
 503 AGACCTTGGCCAGGAGTACGGTGTCTCTTCTGGAGACCGCCAGAGCTGGCATCA 562  
 577 ATGTGAGTATGACCTTTCTGGCCATCGCCAGGAATGAATACCGGGCGGGCATCAGG 636  
 563 ATGTGAGTATGACCTTTCTGGCCATCGCCAGGAATGAATACCGGGCGGGCATCAGG 622  
 637 CGGATGAGCCAGCTCCAGATCCAGACTATGTAGATGCCAGAGAGCGCTCCAGCT 696  
 623 CGGATGAGCCAGCTCCAGATCCAGACTATGTAGATGCCAGAGAGCGCTCCAGCT 682  
 697 GCTGCTCTCTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGATGC 756  
 683 GCTGCTCTCTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGATGC 742  
 757 AGCCTTCCCGCTCCAGGCGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATG 816  
 743 AGCCTTCCCGCTCCAGGCGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATG 802  
 817 GAGGACTCACTGACAGCGCTCTCTAGCAGGAGCTATCTCAACTCTCTACTTGAGTT 876  
 803 GAGGACTCACTGACAGCGCTCTCTAGCAGGAGCTATCTCAACTCTCTACTTGAGTT 862  
 877 COTGGGCTCTCCCGCATCCACAGGAGGGTAAACACTTACTGCTTTATTTTAAATAGTAC 936  
 863 COTGGGCTCTCCCGCATCCACAGGAGGGTAAACACTTACTGCTTTATTTTAAATAGTAC 922  
 937 ATAATTTAATACCAAAAGGCGCTTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 996  
 923 ATAATTTAATACCAAAAGGCGCTTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 982  
 997 CCTTTTGTCTTCTAGGACTTGGGGGCGGCGCTCCCTCTAAGCATAAACAAAGTGGT 1056  
 983 CCTTTTGTCTTCTAGGACTTGGGGGCGGCGCTCCCTCTAAGCATAAACAAAGTGGT 1042  
 1057 GTTGTCTCAGCTCAGCCCGGAGGACACAGATGCACTTTGGGGGTGAGGCGAGTAAATGA 1116  
 1043 GTTGTCTCAGCTCAGCCCGGAGGACACAGATGCACTTTGGGGGTGAGGCGAGTAAATGA 1102  
 1117 CTCATCGCACCT 1130  
 1103 CTCATCGCACCT 1116

BC016615 1700 bp mRNA linear PRI 05-NOV-2001  
 Homo sapiens, Similar to RAB37, member of RAS oncogene family,  
 clone MGC:21391 IMAGE:4520191, mRNA, complete cds.  
 BC016615  
 BC016615.1 GI:16741620  
 MGC.  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1700)  
 Strausberg, R.  
 Direct Submission  
 Submitted (31-OCT-2001) National Institutes of Health, Mammalian

# REMARK COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalob@bcm.tmc.edu](mailto:villalob@bcm.tmc.edu)  
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK plate: 28 Row: j Column: 15  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein.

## FEATURES source

Location/Qualifiers  
 1..1700  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:21391 IMAGE:4520191"  
 /tissue\_type="Prostate, adenocarcinoma."  
 /clone\_lib="NIH\_MGC\_91"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 3..674  
 /codon\_start=1  
 /product="Similar to RAB37, member of RAS oncogene family"  
 /protein\_id="AAH16615.1"  
 /db\_xref="GI:16741621"  
 /translation="MTGPGAVATRDGEAPERSPPSPSYDLTKRMLLIGDTGVGKTC  
 FLIOFKDGLSTGFIATVDFRKNVTVGVKLVQIWDITAGDERFRSVTHAYRD  
 AQALILLYDITNKSDFNIRAWLTHIEYADQDVVILLLGNKADMSRVSLEDEGT  
 LAREVGVPLETSARTGNVLELAFLAIKELKYRAGHQADEPFIQIRYVRSOKRRS  
 CCSFM"

## CDS

BASE COUNT 437 a 444 c 460 g 359 t  
 ORIGIN

Query Match 35.1%; Score 938.4; DB 9; Length 1700;  
 Best Local Similarity 99.9%; Pred. No. 2.2e-240;  
 Matches 939; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1735 AGGCTGGTCTTGAAATCTGAGCTCAAGCAACCTGCCGCGCTCGCCCTCCCAAGTACTG 1794  
 DB 752 AGGCTGGTCTTGAAATCTGAGCTCAAGCAACCTGCCGCGCTCGCCCTCCCAAGTACTG 811  
 QY 1795 GGATTACAGCAGAGGACCATGCCAGGCTAGATGCTTATCCCAATCCTTTGGCA 1854  
 DB 812 GGATTACAGCAGAGGACCATGCCAGGCTAGATGCTTATCCCAATCCTTTGGCA 871  
 QY 1855 GGCATGACGCTCCACAGCGGATTTCTCAACGAGCTGAAGTGTGTAGCCCTCGGTGA 1914  
 DB 872 GGCATGACGCTCCACAGCGGATTTCTCAACGAGCTGAAGTGTGTAGCCCTCGGTGA 931  
 QY 1915 AGAGCCAGATAAGGAGAAATCCCTTTCTAGGTTTGAATGTGTGTGAAAAAAGAGA 1974  
 DB 932 AGAGCCAGATAAGGAGAAATCCCTTTCTAGGTTTGAATGTGTGTGAAAAAAGAGA 991  
 QY 1975 AATCCCTGGCTCTGGAGCTGGTGGGAGACAAAGATTAGCAAACTCCCTCGACATGAT 2034  
 DB 992 AATCCCTGGCTCTGGAGCTGGTGGGAGACAAAGATTAGCAAACTCCCTCGACATGAT 1051  
 QY 2035 CCTTTTGACCCCAAGCTCTGCTCTCCTCCTGACCAACCATCCCTTCTTAACTTCTC 2094  
 DB 1052 CCTTTTGACCCCAAGCTCTGCTCTCCTCCTGACCAACCATCCCTTCTTAACTTCTC 1111

```
QY 2095 AACAGATACAGGGCCCTAAACTGCTTTACCTCCCTCCTACTGAGTCAGGTTAGGTGGT 2154
D 1112 AACAGATACAGGGCCCTAAACTGCTTTACCTCCCTCCTACTGAGTCAGGTTAGGTGGT 1171
QY 2155 GGGAGGTACCCATTTCGGAGTTAAACCAATGCAATATGAGTAAACAAAGTCAATGTGGG 2214
D 1172 GGGAGGTACCCATTTCGGAGTTAAACCAATGCAATATGAGTAAACAAAGTCAATGTGGG 1231
QY 2215 TATGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTCTCCTTGGTCAATATCTCC 2274
D 1232 TATGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTCTCCTTGGTCAATATCTCC 1291
QY 2275 CAAGCTCTGATCCTGCCATGGGAAGTGGACAGGAACATGAGTCAATGACCTGCAGGC 2334
D 1292 CAAGCTCCGATCCTGCCATGGGAAGTGGACAGGAACATGAGTCAATGACCTGCAGGC 1351
QY 2335 ATCTTTACTGAGCTCTGCCGCTTGGAGGGGAGAGGGGAGGAAGATGATGCGTGC 2394
D 1352 ATCTTTACTGAGCTCTGCCGCTTGGAGGGGAGAGGGGAGGAAGATGATGCGTGC 1411
QY 2395 ACATTTCTGAGCTCTGCAATTTCTTCAAGGCAGAAATCTTCTGTCAGCAGTCAGCG 2454
D 1412 ACATTTCTGAGCTCTGCAATTTCTTCAAGGCAGAAATCTTCTGTCAGCAGTCAGCG 1471
QY 2455 GCTCCAGTTTGGGCCCCGATAGGAAGTTCTCCGTGGCTCCTCAGGCAGAGCAGGGAGG 2514
D 1472 GCTCCAGTTTGGGCCCCGATAGGAAGTTCTCCGTGGCTCCTCAGGCAGAGCAGGGAGG 1531
QY 2515 AGGCTGACATTCGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCGAGGAGATCCATCCCAT 2574
D 1532 AGGCTGACATTCGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCGAGGAGATCCATCCCAT 1591
QY 2575 AGACAGCTCTGGGCTCTTTCGATTTGAGTTTTCAGAAATTAACATGTCAGTATTTGGAAA 2634
D 1592 AGACAGCTCTGGGCTCTTTCGATTTGAGTTTTCAGAAATTAACATGTCAGTATTTGGAAA 1651
QY 2635 GCAAAAAA 2674
D 1652 GCAAAAAA 1691
```

```
RESULT 6
LOCUS AX236084 576 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION AX236084.1 GI:15795891
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Meyers, R.A.
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins
JOURNAL Patent: WO 0164887-A 15 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..576
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 141 a 148 c 167 g 120 t
ORIGIN
Query Match 21.5% Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.7e-143;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 138 ATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCATTCATCAAGACGGG 197
D 1 ATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCATTCATCAAGACGGG 60
```

```
QY 198 GCCCTCTGTCGGGAACCTTCATAGCCACCCTCGGCATAGACTTCAGGAACAAGGTGGT 257
D 61 GCCCTCTGTCGGGAACCTTCATAGCCACCCTCGGCATAGACTTCAGGAACAAGGTGGT 120
QY 258 ACTGTGGATGGCTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCAGGAACGGTTC 317
D 121 ACTGTGGATGGCTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCAGGAACGGTTC 180
QY 318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCTTTCCTTCTCTGTATGACATC 377
D 181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCTTTCCTTCTCTGTATGACATC 240
QY 378 ACCAACAATCTTCTTCGACAACATCAGGCGCTCGCTCAGATGATGATGATGATGCC 437
D 241 ACCAACAATCTTCTTCGACAACATCAGGCGCTCGCTCAGATGATGATGATGATGCC 300
QY 438 CAGAGGGACGTGGTGCATCATGCTAGGCAACAGCGGATATGAGCAGCAAGAGTG 497
D 301 CAGAGGGACGTGGTGCATCATGCTAGGCAACAGCGGATATGAGCAGCAAGAGTG 360
QY 498 ATCCGTTCCGAAGACGGAGACACCTTGCCAGGGAGTACGGTGTTCCTTCTGGAGACC 557
D 361 ATCCGTTCCGAAGACGGAGACACCTTGCCAGGGAGTACGGTGTTCCTTCTGGAGACC 420
QY 558 AGCCCAAGACTGGCATGAATGTGAGTACGCTTCTGGCCATCGCCCAAGAACTGAAA 617
D 421 AGCCCAAGACTGGCATGAATGTGAGTACGCTTCTGGCCATCGCCCAAGAACTGAAA 480
QY 618 TACGGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
D 481 TACGGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540
QY 678 CAGAAGAAGCGCTCCAGCTGCTGCTTCATGTGA 713
D 541 CAGAAGAAGCGCTCCAGCTGCTGCTTCATGTGA 576
```

```
RESULT 7
LOCUS AF233582 690 bp mRNA linear ROD 02-MAY-2000
DEFINITION Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
ACCESSION AF233582
VERSION AF233582.1 GI:7677421
KEYWORDS Mus musculus.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 690)
Masuda, E.S., Luo, Y., Young, C., Shen, M., Rossi, A.B., Huang, B.C.,
Yu, S., Bennett, M.K., Payan, D.G. and Scheller, R.H.
TITLE Rab37, is a novel mast cell specific GTPase localized to secretory
granules
JOURNAL FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE 20189834
PUBMED 10722846
REFERENCE 2 (bases 1 to 690)
AUTHORS Luo, Y., Huang, B.C.B., Yu, S., Shen, M. and Masuda, E.S.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
Avenue, South San Francisco, CA 94080, USA
FEATURES Location/Qualifiers
source 1..690
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..690
CDS 1..672
/gene="Rab37"
/codon_start=1
/product="GTPase Rab37"
/protein_id="AAF67162.1"
/db_xref="GI:7677422"
```





```

CDS
/gene="v46133"
/locus="v46133"
/feature="RAS-related protein"
/codon_start=1
/product="RAB-26"
/protein_id="BA84707.1"
/db_xref="GI:5931612"
VKVQLQMDTAGQERFSVTHAYRDAHALLLYDVTNKASFDNIQAWLEIHEYAOH
DVALMLGNKVDSAHVRVKKREDGEKLAKEYLGFPMETSAKTGLNVLDTAIAELK
ORSMKAPSEPRFLHDYVKREGGASCCRP"
BASE COUNT 273 a 379 c 395 g 273 t
ORIGIN
Query Match 11.4%; Score 304.6; DB 9; Length 1320;
Best Local Similarity 71.1%; Pred. No. 2e-70;
Matches 403; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
QY 135 GTGATGCTCTCGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGAC 194
DB 1 GTGATGCTCTCGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGAC 60
QY 195 GGGGCTTCCTCTCGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGAC 254
DB 61 GGTGCTTCCTCTCGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGAC 120
QY 255 GTGATGCTCTCGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGAC 314
DB 121 CTGGACGTGGATGTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 180
QY 315 TTCGAGCGTCACCCATGCTTTACAGAGATGCTCAGGCGTTCGCTGCTGTATGAC 374
DB 181 TTCGAGCGTCACCCATGCTTTACAGAGATGCTCAGGCGTTCGCTGCTGTATGAC 240
QY 375 ATCAACAAATCTCTTCGCAACATCAGGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434
DB 241 CTCACCAAGAGCGCTCTTTGACACATCAGGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 435 GCCAGAGAGCGTGGTGATCATGCTGTAGCAACAGGCGGATATGAGCAGCAAGA 494
DB 301 GCCAGAGAGCGTGGTGATCATGCTGTAGCAACAGGCGGATATGAGCAGCAAGA 360
QY 495 GTGATCGCTTCGAGAGAGAGACCTTGGCCAGGAGTACGCTGCTTCCTTCCTGGAG 554
DB 361 GTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 555 ACCAGCGCAAGACTGGCATCAATGTGGAGTGTAGCTTCTGCGCATCGCCAAAGAACTG 614
DB 421 ACCAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 615 AAATACCGGGCGGGCATFAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674
DB 481 AAGCAGCGCTCCATGAAGGCTCCAGCAGCGCGGCTTCCGGCTGCTGATGATACGTTAAG 540
QY 675 TCCAGAGAGAGCGCTCCAGCTGCTGC 701
DB 541 AGGAGGCTCAGGGGCGCTCCTGCTGC 567

RESULT 11
AF498952 Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete cds.
AF498952 Homo sapiens
AF498952.1 GI:20379079
ACCESSION AF498952
VERSION AF498952.1
KEYWORDS Homo sapiens
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 573)
Puhl, H. L. III, Ikeda, S. R. and Aronstam, R. S.
Homo sapiens RAB family small GTP binding protein RAB26
Unpublished
2 (bases 1 to 573)
Puhl, H. L. III, Ikeda, S. R. and Aronstam, R. S.
Direct Submission
Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
Location/Qualifiers
source
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
1..573
/gene="RAB26"
1..573
/gene="RAB26"
/feature="RAB family member"
/codon_start=1
/product="small GTP binding protein RAB26"
/protein_id="AAM21100.1"
/db_xref="GI:20379080"
/translation="MLVDSGVGKTCLLVRFKDGAFLAGTFISTVGDIFRNKVLVDVG
VKVQLQMDTAGQERFSVTHAYRDAHALLLYDVTNKASFDNIQAWLEIHEYAOH
DVALMLGNKVDSAHVRVKKREDGEKLAKEYLGFPMETSAKTGLNVLDTAIAELK
ORSMKAPSEPRFLHDYVKREGGASCCRP"
BASE COUNT 117 a 151 c 191 g 114 t
ORIGIN
Query Match 11.3%; Score 303.2; DB 9; Length 573;
Best Local Similarity 71.1%; Pred. No. 4e-70;
Matches 401; Conservative 1; Mismatches 163; Indels 0; Gaps 0;
QY 138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGACGG 197
DB 1 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCCAATTCAAAGACGG 60
QY 198 GCCTTCTCTCGGGAGACCTTCATAGCCACCGCTCGGATAGACTTCAGAGAACAGGTGGT 257
DB 61 GCCTTCTCTCGGGAGACCTTCATCTCCACCGTAGGATGACTTCCCGAACAAAGATTCTG 120
QY 258 ACTGTGATGCGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 317
DB 121 GACGTGATGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 180
QY 318 CGAAGCGTACCCATGCTTATACAGAGATGCTCAGGCGCTTCTGCTGCTGCTGCTGCTGCTG 377
DB 181 CGCAGTGTACCCATGCTACTACCGGATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 378 ACCAACAATCTCTTTCGCAACATCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
DB 241 ACCAACAAGGCGCTCTTTGACACATCCAGCGCTGGCTGACCGAGATCCAGAGTACGCC 300
QY 438 CAGAGGAGCGTGGTATCATGCTGTAGGCAACAGGCGGATATCAGCAGGCAAGAGAGTG 497
DB 301 CAGCAGCAGTGGCGCTCATGCTGTGGGAACAGGTTGGAGTCTGCCCATGAGCGTGTG 360
QY 498 ATCCGTTCCCAAGAGAGAGAGACCTTGGCCAGGAGTACGCTGTTTCCCTTCTGCGAGACC 557
DB 361 GTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 558 AGCCCAAGAGTGGCATGATGTGGAGTTAGCTTTCTGGCCATCGCCCAAGAACTGAAA 617
DB 421 AGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 618 TACCGGCGCGGCATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
DB 481 CAGCGTCCATGAAGGCTCCAGCAGGAGCGCGCTTCCGGCTGCTGATGATTTAGAGAGG 540
678 CAGAAGAGCGCTCCAGCTGCTGC 701
541 GAGGTCGAGGGGCGCTCCTGCTGC 564

```

## RESULT 12

AR062279  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AR062279  
Sequence 2 from patent US 5843717.  
AR062279.1 GI:5989970  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 1340)  
Hillman, J.L. and Guegler, K.J.  
Rab protein  
Patent: US 5843717-A 2 01-DEC-1998;  
Location/Qualifiers  
1..1340  
/organism="unknown"  
276 a 387 c 403 g 274 t

Query Match 10.9%; Score 292.6; DB 6; Length 1340;  
Best Local Similarity 70.9%; Pred. No. 3.4e-67;  
Matches 416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY 116 CTAGGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTGGGAAACATGTTT 175  
DB 21 CTAGGACGTGCGCTTCAAGGTGATGCTGTTGGGGGACCTCGGGTGTGGGGAAGACCTGTCT 80  
QY 176 CTGGA-TCCAAATTCAGAGCGGGCTTCTCTGCGGAACCTTCATAGACCCACCTGGGCA 234  
DB 81 GCTGGGTGCGATTCAGAGTGTGCTTCTCTGGGGGACCTTCATCTCCACCGT-AGCA 139  
QY 235 TAGACTTCAGAAACAGGTGGTGACGTGGATGCGGTGAGAGTGAAGCTCAGATCTGGG 294  
DB 140 TTGACTTCCGGAACAAAGTCTCGACGTGATGCTGGAAGTGAAGCTCAGATCTGGG 199  
QY 295 ACACCGCTGGGAGGACGTTTCCGAGCGTCCACCATGCTTATTACAGAGATGCTCAGG 354  
DB 200 ACACAGCTGGTCAAGGAGCGGTTCCGAGTGTATCCCATGCTTACTACCGGGATGCTCATG 259  
QY 355 CTTGCTTCTGCTGTATGACATCACCACAAATCTTCTTCGACAACTCAGGCGCTGGC 414  
DB 260 CTCTGCTGCTCTACGATGTACCAACAGGCGCTCTTTCAGACATCCAGCGCTGGC 319  
QY 415 TCAGTATGATTCAGTATGATCCAGGAGGACGTGTGATCATCTGCTAGGCAACAGG 474  
DB 320 TGACCGAGATCCAGGATGACGCGGACGACGCTGGCGCTCATGCTGCTGGGAACAGG 379  
QY 475 CGGATATGACGAGGAAGAGTATCGTTCCGAAACGAGGAGACCTTGGCCAGGAGT 534  
DB 380 TGGACTCTGCCCATGAGCGGTGGTGAAGAGGAGGAGCGGGGAGAGCTGGCCAAAGGAT 439  
QY 535 ACCGTGTTCTCTTCTGAGACAGCGCCAGAGCTGGCATGATGTGAGCTTACGCTTTC 594  
DB 440 ATGAGTCTGCTTCATGGAGACAGCGCCAGAGCGGCTCAGCTGAGATCTGGCTTCA 499  
QY 595 TGGCCATCGCCAGGAACTGAATACCGGGCGGGCATCAGGCGGATGAGCCAGCTTCC 654  
DB 500 CAGCCATAGCAAGAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCAGCGGAGCGCGCTTC 559  
QY 655 AGATCCGAGATATGTAGATCCCAAGAGAGGCTTCCAGCTGCTGC 701  
DB 560 GGCTGATGATTACGTTAAGAGAGGAGGCTGAGGCGGCTCTCTGCTGC 606

## RESULT 13

AY061826  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AY061826  
Drosophila melanogaster GH21984 full length cDNA.  
AY061826  
AY061826.1 GI:16902019  
FLI\_CDNA.

2536 bp mRNA linear INV 10-NOV-2001

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
BASE COUNT  
ORIGIN

Drosophila melanogaster.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 2536)  
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,  
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,  
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,  
Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.  
Direct Submission  
Submitted (02-NOV-2001) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and  
reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our Web site  
(http://fruitfly.berkeley.edu) or send email to  
cdna@fruitfly.berkeley.edu.

1..2536  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/map="78C9-78D1"  
1..2536  
/gene="CG7605"  
/note="alignment with genomic scaffold AE003594"  
/db\_xref="FLYBASE:FBgn0037072"  
1172..2338  
/gene="CG7605"  
/note="Longest ORF"  
/codon\_start=1  
/product="GH21984p"  
/protein\_id="AAL27637.1"  
/db\_xref="GI:16902020"  
/db\_xref="FLYBASE:FBgn0037072"  
/translation="MASTAVGLGGEGDPCGAGPPAGSAHPDDASSMSDDVFEAETT  
OARIELRRPFPGDSYNPPAAPASVSASITTTTQQQQHHHPSHHSHQSPSH  
HHHHHSQLSLTSGSHYHDDAIPVORSATGPGYRPSREMQMTYADTDDDDNN  
DGNRSYRDEVDMPAPNAHQPPDFDVTNHKTLILGDSGVGKTSFVLYKNTGEFRIG  
SFATVIALTNKVVVDGTRVKLIQIDTAGQERFSVTHAYRDAHALLLLLDVTK  
TTVDNTGALVKEIYEAQEDVIVLIGNKADCSGSRQVRKRGRLGRHNPVFMET  
SANTGLNVELSFTAVAROLKSRGVEHGDGDRFNVDVDRNTARSVCAQCRNN"

831 a 621 c 624 g 460 t

Query Match 8.8%; Score 234.2; DB 3; Length 2536;  
Best Local Similarity 65.7%; Pred. No. 1.8e-51;  
Matches 357; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

QY 131 CAAGGTGATCTCTGGGAGACACAGGCGTGGCGGAAACATGTTTCTGTATCCAAATCAA 190  
DB 1744 CAAGACGATCTCTGTGGGGATTCGGCGGTGGGAAAGACCTCTCTCTGTCCTCAATACAA 1803  
QY 191 AGACGGGCGCTTCTGTGCGGAACTTCATAGCCACCGCTGGCATACACTTCAGGACAA 250  
DB 1804 TACGGCGGAGTTCGGACTGGGTCTCTTTCCGCCACAGATGGGCAATCGCTTACGAAACA 1863



|    |      |   |      |
|----|------|---|------|
| QY | 251  | GGTGGTGACTGTGGATGGCTGGAGATGAAGCTGCAGATCTGGGACACCGCTGGGCAGGA   | 310  |
| Db | 1864 | AGTGGTGGTGGTGGATGGACCGCGTCAAGCTGCAAACTCTGGGACACACGCTGGTCAAGGA | 1923 |
| QY | 311  | ACGGETTCCGAAGCGGTCAACCATGCTATTACAGAGATGCTCAGGCGCTTGGTCTGCTGTA | 370  |
| Db | 1924 | GCGATTCGGAGCGGTTACCCACGCGCTATTATCGGACGCGACGCTCTACTGCTGCTGTA   | 1983 |
| QY | 371  | TGACATCACCAACAATCTTCCTTCGACACATCATGGCGCTGGCTCACTCAGATTCATGA   | 430  |
| Db | 1984 | CGAGCTGACCAACAAGACCACTATGATGACAACTTCGCGCTGGCTGGGCGAGATCCGCGGA | 2043 |
| QY | 431  | GTATGCCAGAGGAGCGTGGTGATCATGCTGCTAGGCAACAAGCGGATATGA---CGAG    | 487  |
| Db | 2044 | GTACGCGCAGGAGGACGTGGTGCTATCGTTTAAATAGGCAACAAGCGGACTGCAGCGCAG  | 2103 |
| QY | 488  | CGAAGAGTGATCCGTTCCGAACGAGGAGACCTTGGCCAGGAGGTACGGTGTTCGCTT     | 547  |
| Db | 2104 | CGAGCGCAGGTGAAGCGGGAGGATGGGAGCGCTTGGGCGGGAGGACCAACATGGCCCTT   | 2163 |
| QY | 548  | CCTGGAGACAGCGCCCAAGACTGGCATGAATGGAGTTAGCGCTTTCGGCCATCGGCAA    | 607  |
| Db | 2164 | CATGGAGACCTCGGCCAAGACGGACTCAATGTGGAGCTGTCTCTCACACGGGTGGCCAG   | 2223 |
| QY | 608  | GGAATCTGAATACCGGGCCGGGCATAGGCGGATGAGCCAGCTTCCAGATCCGAGACTA    | 667  |
| Db | 2224 | GCACTTAAGAGTCGCGGCTACGAGCACGCGCATGATGGAAGGTTCAATGTGCATGATTT   | 2283 |
| QY | 668  | TGT   | 670  |
| Db | 2284 | TGT   | 2286 |

RESULT 14  
HS941F9/C

| LOCUS      | HS941F9   | 127587 bp | DNA | linear | PRI 12-DEC-1999 |
|------------|---|-----------|-----|--------|-----------------|
| DEFINITION | Human DNA sequence from clone CWA-941F9 on chromosome 22q13. Contains the 3' end of the FBLN1 gene for Fibulin 1 isoforms B, C and D, the first exon of the gene for a novel protein (the ortholog of mouse brain protein E46), ESRs, STSs, GSSs and two putative CpG islands, complete sequence. |           |     |        |                 |

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
1 (bases 1 to 127587)  
AUTHORS  
Connor, R.  
Mammalia; Eulheria; Primates; Catarrhini; Hominoidea; Homo.

**TITLE**  
Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone  
requests: clonesrequest@sanger.ac.uk  
**COMMENT**  
On Dec 13, 1999, this sequence version replaced qi:281936.

**COMMENT**

CDS

misc feature

```
/gene="FBLN1"  
/note="match: STS: Em: R76273"
```



```

polyA_signal      complement(10783..10788)
polyA_site        /gene="FBLN1"
polyA_site        10813
polyA_site        /gene="FBLN1"
repeat_region     10819
repeat_region     /note="MIR repeat: matches 109..144 of consensus"
repeat_region     11114..11149
repeat_region     /note="MIR repeat: matches 75..120 of consensus"
repeat_region     11527..11572
repeat_region     /note="L2 repeat: matches 1992..2328 of consensus"
repeat_region     11651..11934
repeat_region     /note="L2 repeat: matches 2454..2750 of consensus"
repeat_region     12001..12270
repeat_region     /note="L2 repeat: matches 2454..2750 of consensus"
misc_feature      12337..13158
misc_feature      /gene="FBLN1"
misc_feature      /note="match: STS: Em:X53742"
misc_feature      <12345..13158
misc_feature      /gene="FBLN1"
misc_feature      /product="bK941F9.2 (Fibulin 1 (isoform B))"
misc_feature      /note="match: cDNAs: Em:AB007943 Em:AF078528 Em:Y14610
misc_feature      Em:AF011357 Em:X53742 Em:AF019887; match: ESTs: Em:R78704
misc_feature      Em:R33281 Em:R32733 Em:H01971 Em:H02060"
misc_feature      /evidence="not experimental"
misc_feature      <12345..12449
misc_feature      /gene="FBLN1"
misc_feature      /note="match: proteins: Sw:P23143"
misc_feature      /codon_start=1
misc_feature      /evidence="not experimental"
misc_feature      /product="bK941F9.2 (Fibulin 1 (isoform B))"
misc_feature      /protein_id="CAB62961.1"
misc_feature      /db_xref="GI:6572284"
misc_feature      /translation="KSKKGRQNTPAGSSKEDCRVLPWKQGLDTHLDA"
misc_feature      12449..12703
misc_feature      /gene="FBLN1"
misc_feature      /note="match: STS: Em:G59664"
misc_feature      12450..12849
misc_feature      /gene="FBLN1"
misc_feature      /note="match: STS: Em:G11168"
misc_feature      complement(12733..13139)
repeat_region     12895..13143
repeat_region     /note="L2 repeat: matches 2419..2705 of consensus"
repeat_region     complement(12974..13280)
repeat_region     /note="match: STS: Em:G08093; match: STS: Em:G08093"
polyA_signal      13138..13143
polyA_signal      /gene="FBLN1"
polyA_signal      13158
polyA_signal      /gene="FBLN1"
repeat_region     13237..13536
repeat_region     /note="AluSp repeat: matches 1..300 of consensus"
repeat_region     13974..14045
repeat_region     /note="MIR repeat: matches 50..122 of consensus"
repeat_region     14352..14607
repeat_region     /note="MIR repeat: matches 2..262 of consensus"
repeat_region     14791..15080
repeat_region     /note="AluSx repeat: matches 1..290 of consensus"
repeat_region     15714..15856
repeat_region     /note="MLT1C repeat: matches 300..451 of consensus"
repeat_region     16332..16625
repeat_region     /note="MLT1C repeat: matches 1..310 of consensus"
repeat_region     16637..16815
repeat_region     /note="MLT1B repeat: matches 1..190 of consensus"
repeat_region     16816..17125
repeat_region     /note="AluSg repeat: matches 1..310 of consensus"
repeat_region     17126..17338
repeat_region     /note="MLT1B repeat: matches 190..381 of consensus"
repeat_region     17506..17806
repeat_region     /note="AluYb8 repeat: matches 1..318 of consensus"
repeat_region     17807..18030
repeat_region     /note="7 copies 32 mer 64 conserved"
repeat_region     17808..18031
repeat_region     /note="112 copies 2 mer aa 62 conserved"
repeat_region     18046..18303

```

---

```

repeat_region     /note="HAL1 repeat: matches 351..597 of consensus"
repeat_region     18304..18597
repeat_region     /note="AluSg repeat: matches 1..309 of consensus"
repeat_region     18598..18822
repeat_region     /note="HAL1 repeat: matches 597..825 of consensus"
repeat_region     19070..19305
repeat_region     /note="L1MC5 repeat: matches 6484..6730 of consensus"
repeat_region     19306..19621
repeat_region     /note="AluSg repeat: matches 1..310 of consensus"
repeat_region     19622..20401
repeat_region     /note="L1MC5 repeat: matches 6730..7471 of consensus"
repeat_region     20402..20711
repeat_region     /note="AluSx repeat: matches 1..312 of consensus"
repeat_region     20712..20770
repeat_region     /note="L1MC5 repeat: matches 7471..7531 of consensus"
repeat_region     20851..21158
repeat_region     /note="AluSg repeat: matches 1..301 of consensus"

```

---

```

Query Match      8.08; Score 214.8; DB 9; Length 127587;
Best Local Similarity 80.3%; Pred. No. 7.5e-46;
Matches 252; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy      1526 TTTGTTTTTATTTTTTGAATGAGTCTGTTCTGTGCGCCAGGCTGAGTGCAGTAGT 1585
Db      59011 TTTGTTTTTCTTTTTTGAACGGAGTCTACTCTGCACCCAGGCTGGAGTGCATATTTT 58952
Qy      1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCTCCACCTCAGC 1645
Db      58951 GCGATCTCGGCTTATTGAGCTCTGCTCCCGGATTCAGTGTCTCTCTGCTCCTCAGC 58892
Qy      1646 CGCGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGCTAAATTTTGTATTTTT 1705
Db      58991 CTCCCAAGTAGCTGGAATTACAGGTGCATACACACCTGGCTAAATTTTATTTT 58832
Qy      1706 TGTAGACACAGGTTTTCGCCATGTTGCCAGCTGTCTTGAATTCCTGAGCTCAAGCAA 1765
Db      58831 AGTAGAGATGGGTTTCACCATGTTGCCAGCTGTCTTGAATTCCTGAGCTCAAGTGA 58772
Qy      1766 CTGCGCGCTCGGCCCTCCCAAGTACTGGGATTACAGCAGAGGACCATGCCAGGC 1825
Db      58771 TCTGCTGCTCGGCCCTCCCAAGTGTCCGAGTACAGGCAAGACCATGCCGCGCC 58712
Qy      1826 TAGATGTGCTTAT 1839
Db      58711 TGGTTTCTGTGTTT 58698

```

---

```

RESULT 15
AC010352/c      AC010352      102332 bp      DNA      linear      PRI 18-APR-2000
LOCUS          Homo sapiens chromosome 5 clone CTD-2022F20, complete sequence.
DEFINITION     AC010352
ACCESSION      AC010352
VERSION        AC010352.4      GI:7109394
KEYWORDS       HTG.
SOURCE         Homo sapiens.
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 102332)
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL        Direct Submission
REFERENCE      2 (bases 1 to 102332)
AUTHORS       DOE Joint Genome Institute.
TITLE         Direct Submission
JOURNAL        Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
REFERENCE      3 (bases 1 to 102332)
AUTHORS       Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE         DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL        Submitted (29-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
REFERENCE      4 (bases 1 to 102332)
AUTHORS       Drive! Walnut Creek, CA 94598, USA
TITLE         DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL        Submitted (29-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
REFERENCE      4 (bases 1 to 102332)

```

|            |   |
|------------|---|
| AUTHORS    | DOE Joint Genome Institute and Stanford Human Genome Center.  |
| TITLE      | Direct Submission   |
| JOURNAL    | Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  |
| COMMENT    | On Feb 29, 2000 this sequence version replaced gi:6606056.<br>Draft Sequence Produced by DOE Joint Genome Institute<br>www.jgi.doe.gov<br>Finishing Completed at Stanford Human Genome Center<br>www.shgc.stanford.edu<br>Quality: Phrap Quality >=40 99.4% of Sequence;<br>Estimated Total Number of Errors is 1.5.<br>STS Content:<br>WI-12799 G24393<br>SHGC-13554 G14788. |
| FEATURES   | Location/Qualifiers<br>source 1. .102332<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/chromosome="5"<br>/clone="CTD-2022F20"  |
| BASE COUNT | 29578 a 22934 c 23227 g 26593 t   |
| ORIGIN     |   |
|            | Query Match 8.0%; Score 214.2; DB 9; Length 102332;<br>Best Local Similarity 82.6%; Pred. No. Ie=45;<br>Matches 257; Conservative 0; Mismatches 53; Indels 1; Gaps 1;   |
| QY         | 1522 GAGTTTGTGTTTATTTTTTGAATGAGTCTCGTTCTGTCGCCAGGCTGAGGTGCAG 1581   |
| Db         |   |
|            | 22554 GATTTTATTTATTTTTTTTGAGACACAGTCTCGCTCTGTCGCCAGGCTGGAATGCAG 22495   |
| QY         | 1582 TAGTGCAATCTCGGCTCACTCAACCTCCACTCCCTGGGGCTCAAGCATCTCTCCACCT 1641  |
| Db         | TGGTGCATCTCGGCTCCCTGCAACCTCCACTTCCT -GGGTGAAGCATTTCTCTCGCT 22436  |
| QY         | 1642 CAGCCGCCGAAGTAGTGCGGACTATAGTGTGTACCATCACACCTGGCTAATTTTGTAT 1701  |
| Db         |   |
|            | 22435 CAGCCTCCGTGACTAGCTGGGACTACAGGGGTGTGCCACCACACCCGGCTAATTTTGTAT 22376  |
| QY         | 1702 TTTTGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGTAATTCCTGAGCTCAA 1761   |
| Db         |   |
|            | 22375 TTTTGTAGTAGAGTGGGGTTTCCACCATGTTGCCAGGCTGGTCTGGAATCCTCGACCTCAG 22316   |
| QY         | 1762 GCAACTCGCGGCTCTGGGCTCCCAAAGTACTGGGATTACACGAGAGGACCATGCC 1821   |
| Db         |   |
|            | 22315 GCATCTGCCATCTCGGCTCCCAAAGTGTGGGATTACAGCATGAGCCCATGCC 22256  |
| QY         | 1822 AGGCTAGATGT 1832   |
| Db         |   |
|            | 22255 AGACTATTTAT 22245   |

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 20:33:28 ; Search time 568 Seconds  
(without alignments)  
10601.837 Million cell updates/sec

Title: US-09-817-199b-1

Perfect score: 2674  
Sequence: 1 ttgcctgctggccgacct.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125995159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 2581.6 | 96.5        | 2623   | 22 | AAS27036           |
| 2          | 2240   | 83.8        | 2323   | 22 | CDNA encoding nove |
| 3          | 1286   | 48.1        | 1316   | 22 | Human polynucleoti |
| 4          | 1099.6 | 41.1        | 1106   | 22 | CDNA encoding nove |
| 5          | 1079   | 40.4        | 1109   | 22 | Human polynucleoti |
| 6          | 873    | 32.6        | 875    | 22 | Nucleotide sequenc |
| 7          | 795.4  | 29.7        | 797    | 21 | Human Ras protein  |
| 8          | 568    | 21.2        | 576    | 22 | Eosinophil activat |
| 9          | 550.8  | 20.6        | 543    | 20 | Human degranulatio |
|            |        |             |        |    | DNA sequence encod |

|    |       |      |        |    |          |                    |
|----|-------|------|--------|----|----------|--------------------|
| 10 | 480   | 18.0 | 576    | 24 | ABA02773 | Mouse degranulatio |
| 11 | 468.6 | 17.5 | 475    | 23 | AAS81431 | DNA encoding novel |
| 12 | 292.6 | 10.9 | 1340   | 19 | AAV65197 | Human RAB protein, |
| 13 | 274   | 10.2 | 443    | 21 | AA89717  | Mouse Exo104 nucle |
| 14 | 237   | 8.9  | 2001   | 23 | ABL29523 | Drosophila melanog |
| 15 | 209.4 | 7.8  | 964    | 22 | AAS27040 | CDNA encoding nove |
| 16 | 207.4 | 7.8  | 24000  | 21 | AA88551  | Human dual-specifi |
| 17 | 206.6 | 7.7  | 1786   | 22 | AAD08354 | Human secreted pro |
| 18 | 204.8 | 7.7  | 10859  | 22 | AAJ04720 | Human reproductive |
| 19 | 204.8 | 7.7  | 10859  | 23 | ABL97627 | Human testicular a |
| 20 | 203.4 | 7.6  | 138169 | 21 | AAA34791 | Human adenosine re |
| 21 | 203.4 | 7.6  | 141589 | 21 | AAF20913 | Human ELAM-1 polyn |
| 22 | 203.4 | 7.6  | 141589 | 21 | AAF21127 | Human low adenosin |
| 23 | 203.4 | 7.6  | 141589 | 21 | AAF21152 | Human adenosine re |
| 24 | 203.4 | 7.6  | 141589 | 21 | AAA35005 | Human adenosine re |
| 25 | 203.4 | 7.6  | 141589 | 21 | AAA35030 | Human adenosine re |
| 26 | 203.4 | 7.6  | 146981 | 21 | AAF21442 | Human ELAM-1 polyn |
| 27 | 203.4 | 7.6  | 209273 | 21 | AAF21437 | Human factor-relat |
| 28 | 203   | 7.6  | 409    | 22 | AAK70203 | Human immune/haema |
| 29 | 202.8 | 7.6  | 27579  | 22 | AAK66484 | Human immune/haema |
| 30 | 202.6 | 7.6  | 29222  | 24 | ABL39412 | Human electron-tra |
| 31 | 202.4 | 7.6  | 10468  | 22 | AAK77961 | Human immune/haema |
| 32 | 201.8 | 7.5  | 103747 | 24 | ABQ88139 | Human osteoblast d |
| 33 | 201.6 | 7.5  | 52845  | 22 | AAK71437 | Human immune/haema |
| 34 | 201.4 | 7.5  | 4969   | 22 | ABA16169 | Human nervous syst |
| 35 | 201.4 | 7.5  | 6720   | 22 | ABQ88138 | Human osteoblast d |
| 36 | 201.2 | 7.5  | 33353  | 22 | AAK70003 | Human immune/haema |
| 37 | 201   | 7.5  | 7428   | 22 | ABA07164 | Human pancreatic c |
| 38 | 200.8 | 7.5  | 21835  | 22 | AAK66483 | Human immune/haema |
| 39 | 200.6 | 7.5  | 21833  | 22 | AAL35910 | Human musculoskele |
| 40 | 200.4 | 7.5  | 10663  | 22 | ABA18961 | Human nervous syst |
| 41 | 200.4 | 7.5  | 31865  | 22 | AAK85330 | Human immune/haema |
| 42 | 200.4 | 7.5  | 31865  | 22 | AAK85331 | Human immune/haema |
| 43 | 200.2 | 7.5  | 21777  | 22 | AAK74096 | Human immune/haema |
| 44 | 199.8 | 7.5  | 1130   | 22 | ABA20629 | Human nervous syst |
| 45 | 199.8 | 7.5  | 1130   | 22 | AAL37008 | Human musculoskele |

#### ALIGNMENTS

RESULT 1

AAS27036

ID AAS27036 standard; CDNA; 2623 BP.

XX AC

XX AAS27036;

XX 07-NOV-2001 (first entry)

XX DE

XX CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
 KW organ transplant rejection; infection; hepatitis C; blood disorder;  
 KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
 KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
 KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;  
 KW acquired immune deficiency syndrome.

XX OS

XX Homo sapiens.

XX PN

XX WO200154733-A1.

XX PD

XX 02-AUG-2001.

XX PF

XX 17-JAN-2001; '2001WO-US01312.

XX XX

XX 31-JAN-2000; '2000US-0179065.

XX PR

XX 04-FEB-2000; '2000US-0180628.

XX PR

XX 24-FEB-2000; '2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465460/50.

P-PSDB; AAU17119.

Novel polypeptides useful for diagnosing, treating, preventing and/or  
prognosing disorders related to the proteins, including cancers, immune  
disorders and neuronal disorders

Claim 1; SEQ ID No 71; 880pp; English.

XX The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, hemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders, in  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
 CC pathway protein coding sequences and PCR primers of the invention.  
 XX

Query Match 96.5%; Score 2581.6; DB 22; Length 2623;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
 42 ATGACGGGACGCGCGGCGGTGCGACCGCGGATGGCGAGCCCGCGAGCGCTCCCG 101  
 12 ATGACGGGACGCGCGCGGCGGTGCGACCGCGGATGGCGAGCCCGCGAGCGCTCCCG 71  
 102 CCTGCGAGTCCGAGCTACGACCTACCGGCAAGGTGATGCTTCTGGGAGACACAGCGTC 161  
 72 CCTGCGAGTCCGAGCTACGACCTACCGGCAAGGTGATGCTTCTGGGAGACACAGCGTC 131  
 162 GGCAGAACATGTTCTCTGATCCCAATTCAGAGCGGGCGCTTCTGTCGGAACCTTCATA 221  
 132 GGCAGAACATGTTCTCTGATCCCAATTCAGAGCGGGCGCTTCTGTCGGAACCTTCATA 191  
 222 GCCACCGTGGCATAGACTTCAGGAACAAGGTGGTGACTGGATGGCGTGAGAGTGAAG 281  
 192 GCCACCGTGGCATAGACTTCAGGAACAAGGTGGTGACTGGATGGCGTGAGAGTGAAG 251  
 282 CTGACAGATGGGACACCGCTGGCGAGAACGGTTCGGAAGCTCACCCATGCTTATATAC 341  
 252 CTGACAGATGGGACACCGCTGGCGAGAACGGTTCGGAAGCTCACCCATGCTTATATAC 311  
 342 AGAGATGCTCAGGCGCTTCTGCTGTATGATGATCACCAGCAAAATCTTCTTCGACAC 401  
 312 AGAGATGCTCAGGCGCTTCTGCTGTATGATGATCACCAGCAAAATCTTCTTCGACAC 371  
 402 ATCAGGGCGCTGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 461  
 372 ATCAGGGCGCTGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 431  
 462 CTAGGCAACAAGCGGATATGACGAGCAAGAGTATGATGATGATGATGATGATGATGATG 521  
 432 CTAGGCAACAAGCGGATATGACGAGCAAGAGTATGATGATGATGATGATGATGATGATG 491  
 522 TTGGCAGGAGTACGGTGTTCCTTCTGAGACCGACCGCCAGACTGGCATGAATGTG 581  
 492 TTGGCAGGAGTACGGTGTTCCTTCTGAGACCGACCGCCAGACTGGCATGAATGTG 551  
 582 GAGTATGCTTCTGCGCATTCGCAAGGAATACCGGGCGCGGATCAGGGCGAT 641  
 552 GAGTATGCTTCTGCGCATTCGCAAGGAATACCGGGCGCGGATCAGGGCGAT 611  
 642 GAGCCAGCTTCCAGATCCGAGACTATGTAGATCCGAGAAAGCGCTCCAGCTGCTGC 701  
 612 GAGCCAGCTTCCAGATCCGAGACTATGTAGATCCGAGAAAGCGCTCCAGCTGCTGC 671

QY 702 TCCTTCATGTGAATCCACAGGGGACAGAGAGGCTCTGAGGACACAGGATGACGCCT 761  
 DB 672 TCCTTCATGTGAATCCACAGGGGACAGAGAGGCTCTGAGGACACAGGATGACGCCT 731  
 QY 762 TCCCTCCTCCAGCGCTGGCTTATTCGAAGAGGCTGAGCCAAATGGGGAAGAAGATGGAGGA 821  
 DB 732 TCCCTCCTCCAGCGCTGGCTTATTCGAAGAGGCTGAGCCAAATGGGGAAGAAGATGGAGGA 791  
 QY 822 CTCACATGACACGCGCTTCTAGCAGGAGCTATATCTCCAACTCTCTACTTCTAGTTCCTGC 881  
 DB 792 CTCACATGACACGCGCTTCTAGCAGGAGCTATATCTCCAACTCTCTACTTCTAGTTCCTGC 851  
 QY 882 GGTCTCTCCCGCATCCACAGGAGGCTTAAACACTTAGCTTTTATTTAATAGTACATAAT 941  
 DB 852 GGTCTCTCCCGCATCCACAGGAGGCTTAAACACTTAGCTTTTATTTAATAGTACATAAT 911  
 QY 942 TTAATACCAAAAAGCGCTGATCCCAAAAACCGAGCTGGGAGCTAGTGGCGCTT 1001  
 DB 912 TTAATACCAAAAAGCGCTGATCCCAAAAACCGAGCTGGGAGCTAGTGGCGCTT 971  
 QY 1002 TTGCTTTCTAGGACTTTGGGGGCGGCGCTTCCCTCTTAAGCATAAACAAGGTGGTGTGC 1061  
 DB 972 TTGCTTTCTAGGACTTTGGGGGCGGCGCTTCCCTCTTAAGCATAAACAAGGTGGTGTGC 1031  
 QY 1062 TCCAGCTCAGCCCCAGGGGACACAGATGCACCTTTGGGGGTGAGGGCAGGTAAAGTACCA 1121  
 DB 1032 TCCAGCTCAGCCCCAGGGGACACAGATGCACCTTTGGGGGTGAGGGCAGGTAAAGTACCA 1091  
 QY 1122 TCCGACCTCAGTTTCTAGCTGGACAGAGGCTCAGGTGACCCAGCTTCTACTGTCTCCCGC 1181  
 DB 1092 TCCGACCTCAGTTTCTAGCTGGACAGAGGCTCAGGTGACCCAGCTTCTACTGTCTCCCGC 1151  
 QY 1182 TCTCAGGAGCTTATCTTCCGCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAA 1241  
 DB 1152 TCTCAGGAGCTTATCTTCCGCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAA 1211  
 QY 1242 GACCAAGCTCAGGGAAGATAGAGATATGGAGTGGGAGGGGAGGAGCAAGGGGAGA 1301  
 DB 1212 GACCAAGCTCAGGGAAGATAGAGATATGGAGTGGGAGGGGAGGAGCAAGGGGAGA 1271  
 QY 1302 GAGTAGGCTTAGCTGGCTATCTTGGCTTACTTAACA-CCCCCTCGAGGCAATGCCCT 1360  
 DB 1272 GAGTAGGCTTAGCTGGCTATCTTGGCTTACTTAACA-CCCCCTCGAGGCAATGCCCT 1331  
 QY 1361 TTTCTCCAGCACACAGCATTGGGGACCTGGAATAATTTGTTCCAGGCTCTGTCT 1420  
 DB 1332 TTTCTCCAGCACACAGCATTGGGGACCTGGAATAATTTGTTCCAGGCTCTGTCT 1391  
 QY 1421 CTGCACTTCAGATCTCGGGGAGCGCTTCCGCCCTGAATCCCTGCTTAGCTACCTTC 1480  
 DB 1392 CTGCACTTCAGATCTCGGGGAGCGCTTCCGCCCTGAATCCCTGCTTAGCTACCTTC 1451  
 QY 1481 CTGCTCTGCACCTTAAACCTCAGGTGAGAACTAGGAAAGAGTGTGTTTTTATTTTTT 1540  
 DB 1452 CTGCTCTGCACCTTAAACCTCAGGTGAGAACTAGGAAAGAGTGTGTTTTTATTTTTT 1511  
 QY 1541 TTGAATGAGTCTGCTTCTGCTCCAGGCTGAGTGCAGTGTGCAATCTCCGCTCAC 1600  
 DB 1512 TTGAATGAGTCTGCTTCTGCTCCAGGCTGAGTGCAGTGTGCAATCTCCGCTCAC 1571  
 QY 1601 TACAACCTCCACCTCGGGCTCAAGGATCTCCACCTCAGCCCGCCAGTAGCTGG 1660  
 DB 1572 TACAACCTCCACCTCGGGCTCAAGGATCTCCACCTCAGCCCGCCAGTAGCTGG 1631  
 QY 1661 GACTATAGTGTGTGTACCATCACACTGGCTTAAATTTTTTGTATTTTTTGTAGACACAGGTT 1720  
 DB 1632 GACTATAGTGTGTGTACCATCACACTGGCTTAAATTTTTTGTATTTTTTGTAGACACAGGTT 1691  
 QY 1721 TCGCCATGTTGCCAGGCTGGTGTGAAATTTCTGAGCTCAAGCAACTGCCGGCTCGGC 1780  
 DB 1692 TCGCCATGTTGCCAGGCTGGTGTGAAATTTCTGAGCTCAAGCAACTGCCGGCTCGGC 1751  
 QY 1781 CTCCCAAGTACTGGGATTACAGCAGAAAGCCACTGCCAGGCTAGATGTCTTATC 1840



QY 255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGGAACGG 314  
DB |||||  
241 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGGAACGG 300  
QY |||||  
315 TTCCGAAGCGTCACCCATGCTTATTTACAGAGATGCTCAGGCCCTGCTTCTGCTGTATGAC 374  
DB |||||  
301 TTCCGAAGCGTCACCCATGCTTATTTACAGAGATGCTCAGGCCCTGCTTCTGCTGTATGAC 360  
QY |||||  
375 ATCACCACAAATCTTTTCGACACATCAGGGCTGGCTCAGTGAATTCATGAGTAT 434  
DB |||||  
361 ATCACCACAAATCTTTTCGACACATCAGGGCTGGCTCAGTGAATTCATGAGTAT 420  
QY |||||  
435 GCCCAGAGGAGCTGTGATCATGCTGTAGGACACAGCGCGGATATGAGCAGCGAAGA 494  
DB |||||  
421 GCCCAGAGGAGCTGTGATCATGCTGTAGGACACAGCGCGGATATGAGCAGCGAAGA 480  
QY |||||  
495 GTGATCCGTTCCGAAGACGAGAGACCTTTGGCCAGGGAGTACGGTGTCTCCCTTCCTGGAG 554  
DB |||||  
481 GTGATCCGTTCCGAAGACGAGAGACCTTTGGCCAGGGAGTACGGTGTCTCCCTTCCTGGAG 540  
QY |||||  
555 ACCAGCGCAGACTGGCATGATGTGGAGTTAGCCTTTCTGGCCATCGCCCAAGGAACATG 614  
DB |||||  
541 ACCAGCGCAGACTGGCATGATGTGGAGTTAGCCTTTCTGGCCATCGCCCAAGGAACATG 600  
QY |||||  
615 AAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674  
DB |||||  
601 AAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 660  
QY |||||  
675 TCCAGAGAAGCGTCCAGCTGCTCTCTTCATGTGAATCCAGGGGCGAGAGGAG 734  
DB |||||  
661 TCCAGAGAAGCGTCCAGCTGCTCTCTTCATGTGAATCCAGGGGCGAGAGGAG 720  
QY |||||  
735 GCTCTGGAGGCACAGGATGAGCAGCTTCCCGCTCCAGGCGCTGGCTTATTTCCAGAGGC 794  
DB |||||  
721 GCTCTGGAGGCACAGGATGAGCAGCTTCCCGCTCCAGGCGCTGGCTTATTTCCAGAGGC 780  
QY |||||  
795 TGAGCCATGGGGAGGAAGATGGAGACTCACTGCACAGCCGCTTCCCTAGCAGGGAGCTA 854  
DB |||||  
781 TGAGCCATGGGGAGGAAGATGGAGACTCACTGCACAGCCGCTTCCCTAGCAGGGAGCTA 840  
QY |||||  
855 TACTCAACTCTACTTTGAGTTCTCGCGTCTCCCGCTCCAGGAGGCTGTAACACAC 914  
DB |||||  
841 TACTCAACTCTACTTTGAGTTCTCGCGTCTCCCGCTCCAGGAGGCTGTAACACAC 900  
QY |||||  
915 TTAGCTTTTATTTTATAGTACATATTTAATACCAAAAAGCGCCTGGATCCCCAAA 974  
DB |||||  
901 TTAGCTTTTATTTTATAGTACATATTTAATACCAAAAAGCGCCTGGATCCCCAAA 960  
QY |||||  
975 AACCGAGGCTGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTTGGGGGCGCGCCCTCCC 1034  
DB |||||  
961 AACCGAGGCTGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTTGGGGGCGCGCCCTCCC 1020  
QY |||||  
1035 TCCTAAGCATAAAGGTGGTGTCTCCAGCTCAGCCCGCCAGGGGACACAGATGCATTT 1094  
DB |||||  
1021 TCCTAAGCATAAAGGTGGTGTCTCCAGCTCAGCCCGCCAGGGGACACAGATGCATTT 1080  
QY |||||  
1095 TGGGGGTAGGCGAGTAATGACTCCATCGACCCCTCAGTTTCAGCTGGAGAGGCTCAG 1154  
DB |||||  
1081 TGGGGGTAGGCGAGTAATGACTCCATCGACCCCTCAGTTTCAGCTGGAGAGGCTCAG 1140  
QY |||||  
1155 GTGACCCAGCCTTCACTGCTCCCGCTCTCAGGAGCTTATCTTCGCCCCCAATCTCCCAA 1214  
DB |||||  
1141 GTGACCCAGCCTTCACTGCTCCCGCTCTCAGGAGCTTATCTTCGCCCCCAATCTCCCAA 1200  
QY |||||  
1215 ATAAGTGGCCCTTGTGCTGTGAGGAAGACCAAGCCTCAGGAAGATAGAGATGGA 1274  
DB |||||  
1201 ATAAGTGGCCCTTGTGCTGTGAGGAAGACCAAGCCTCAGGAAGATAGAGATGGA 1260  
QY |||||  
1275 GATGGAGGGGAGGACAAAGGGCGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCCTTAC 1334  
DB |||||  
1261 GATGGAGGGGAGGACAAAGGGCGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCCTTAC 1320

RESULT 3  
AAS27458  
ID AAS27458 standard; cdna; 1316 BP.  
XX

QY 1335 TAAACACCCCTGGAGGCAATGCCCTTTTCTCCAGACACAGCAATTTGGGGCAGCTGG 1394  
DB |||||  
1321 TAAACACCCCTGGAGGCAATGCCCTTTTCTCCAGACACAGCAATTTGGGGCAGCTGG 1380  
QY |||||  
1395 AAATATTGTTCCAGGCTCCTGTTCTCGACTTCAGATCTCCGGGAGGCCCTCCCGCC 1454  
DB |||||  
1381 AAATATTGTTCCAGGCTCCTGTTCTCGACTTCAGATCTCCGGGAGGCCCTCCCGCC 1440  
QY |||||  
1455 CTTGAATCCTGGCTTAGCTACCTTCTGCCTGTGCACCTTAAAACTCAGSTCAGAACT 1514  
DB |||||  
1441 CTTGAATCCTGGCTTAGCTACCTTCTGCCTGTGCACCTTAAAACTCAGSTCAGAACT 1500  
QY |||||  
1515 AGGAAAGAGTTTGTATTTTATTTTGAATGGAGTCTGTTCTGTCGCCAGGCTGA 1574  
DB |||||  
1501 AGGAAAGAGTTTGTATTTTATTTTGAATGGAGTCTGTTCTGTCGCCAGGCTGA 1560  
QY |||||  
1575 GGTGCAGTAGTGAATCTCGGCTCACTACACCTCCACTCCCTGGGGCTCAACGATCCT 1634  
DB |||||  
1561 GGTGCAGTAGTGAATCTCGGCTCACTACACCTCCACTCCCTGGGGCTCAACGATCCT 1620  
QY |||||  
1635 CCCACTCAGCCCGAAGTAGCTGGGACTATAGTGTGTACCATCACACTGGCTTAAT 1694  
DB |||||  
1621 CCCACTCAGCCCGAAGTAGCTGGGACTATAGTGTGTACCATCACACTGGCTTAAT 1680  
QY |||||  
1695 TTTGTATTTTGTAGACACAGGTTTGGCCATGTTGCCAGGCTGGTCTTGAATTCCTG 1754  
DB |||||  
1681 TTTGTATTTTGTAGACACAGGTTTGGCCATGTTGCCAGGCTGGTCTTGAATTCCTG 1740  
QY |||||  
1755 AGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTACTGGGATTTACAGCAGAGGCAC 1814  
DB |||||  
1741 AGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTACTGGGATTTACAGCAGAGGCAC 1800  
QY |||||  
1815 CATGCCAGGCTAGATGTGTTATCCCAATCCTTTGGCAGGCTCAGCTCCACAGGCG 1874  
DB |||||  
1801 CATGCCAGGCTAGATGTGTTATCCCAATCCTTTGGCAGGCTCAGCTCCACAGGCG 1860  
QY |||||  
1875 ATTTCCTCAAGCAGCTGAAGTGTGTAGCCCTCTCGGGTTAAGAGCAGATGAAGAGAAAT 1934  
DB |||||  
1861 ATTTCCTCAAGCAGCTGAAGTGTGTAGCCCTCTCGGGTTAAGAGCAGATGAAGAGAAAT 1920  
QY |||||  
1935 CCCTTTCTTAGGTTTGAATGTGTTGTAATAAAGAGAAATCCCTGGCTCCTGGAGCT 1994  
DB |||||  
1921 CCCTTTCTTAGGTTTGAATGTGTTGTAATAAAGAGAAATCCCTGGCTCCTGGAGCT 1980  
QY |||||  
1995 GTGGGAGACAGATTAAAGCAACCTCCCTCGATGTATPCCTTTTGACCCCAAGCTCTG 2054  
DB |||||  
1981 GTGGGAGACAGATTAAAGCAACCTCCCTCGATGTATPCCTTTTGACCCCAAGCTCTG 2040  
QY |||||  
2055 CTTCTCTCCCTGACCCACCCATGCCCTTTCTTAACTTCTCAAAACAGATACCAGGGCTAA 2114  
DB |||||  
2041 CTTCTCTCCCTGACCCACCCATGCCCTTTCTTAACTTCTCAAAACAGATACCAGGGCTAA 2100  
QY |||||  
2115 ACTCTCTTACCTCCCTCTACTGAGTTCAGTTAGTGGTGGGAGG---TCACCCATTTTC 2171  
DB |||||  
2101 ACTCTCTTACCTCCCTCTACTGAGTTCAGTTAGTGGTGGGAGGTTCAACCCATTTTC 2160  
QY |||||  
2172 CGAGTT--AAACCAATGCAATA--TGAGTAAACAAAGT---GAGTGGGTATCTCTGGG 2225  
DB |||||  
2161 CGAGTTAAACCCCAATGCAATATTGAGTAAACAAAGTTCATGTGGGTATTCTCTGGG 2220  
QY |||||  
2226 TAGAGAGAGGGTAGCAAGTTCAATGTCTCTCTGTCATATCTCCCAAGCTCTGA 2285  
DB |||||  
2221 TAGAGAGAGGGTAGCAAGTTCAATGTCTCTCTGTCATATCTCCCAAGCTCTGA 2280  
QY |||||  
2286 TCCTGCCATGGGAAGTGGACAGGAACATGAGTGCATGACCT 2328  
DB |||||  
2281 TCCTGCCATGGGAAGTGGACAGGAACATGAGTGCATGACCT 2323





PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI P-PSDB; AAU17541.  
 XX  
 DR WPI: 2001-465460/50.  
 DR P-PSDB; AAU17541.  
 XX  
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders -  
 XX  
 PS Claim 1: SEQ ID No 493; 880pp; English.  
 XX  
 CC The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders, cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
 CC pathway protein coding sequences and PCR primers of the invention.  
 XX

Query Match 48.1%; Score 1286; DB 22; Length 1316;  
 Best Local Similarity 98.5%; Pred. No. 5.6e-237;  
 Matches 1286; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
 45 ACGGGCAGCCAGGCGCGCTTCCACCCCGGGATGGCGAGGCCCGCCAGCGCTCCCGCGCC 104  
 1 ACGGGCAGCCAGGCGCGCTTCCACCCCGGGATGGCGAGGCCCGCCAGCGCTCCCGCGCC 60  
 105 TGCAGTCCGAGCTACGACCTACAGGCAAGGTGATGCTTCTGGGACACACAGCGCTCGGC 164  
 61 TGCAGTCCGAGCTACGACCTACAGGCAAGGTGATGCTTCTGGGACACACAGCGCTCGGC 120  
 165 AAAACATGTTCTGATCCAAATCAAGACGGGCGCTTCTGTCGGAAACCTTCATAGCC 224  
 121 AAAACATGTTCTGATCCAAATCAAGACGGGCGCTTCTGTCGGAAACCTTCATAGCC 180  
 225 ACCGTGGCATAGACTTCAGGAACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 284  
 181 ACCGTGGCATAGACTTCAGGAACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240

QY 285 CAGATCTGGGACACCCTGGGAGGAAACGGTTCGGAAGCGTCAACCATGCTTATTACAGA 344  
 DB 241 CAGATCTGGGACACCCTGGGAGGAAACGGTTCGGAAGCGTCAACCATGCTTATTACAGA 300  
 QY 345 GATGCTCAGGCGCTTCTGCTGTATGACATCACCACAAATCTTCTTTCGACACATC 404  
 DB 301 GATGCTCAGGCGCTTCTGCTGTATGACATCACCACAAATCTTCTTTCGACACATC 360  
 QY 405 AGGGCCTGCTCACTGAGATTTCATGAGTATGCCAGAGGACGCTGTGATCATGCTGCTA 464  
 DB 361 AGGGCCTGCTCACTGAGATTTCATGAGTATGCCAGAGGACGCTGTGATCATGCTGCTA 420  
 QY 465 GGCACAAGCGGATATGAGCAGCGAAAGATGATCCGTTCCGAAGACGGAGACCTTG 524  
 DB 421 GGCACAAGCGGATATGAGCAGCGAAAGATGATCCGTTCCGAAGACGGAGACCTTG 480  
 QY 525 GCCAGGAGTACGGTCTTCCCTTCTGGAGACCGCCAGAGCTGGCATGATGTTGGAG 584  
 DB 481 GCCAGGAGTACGGTCTTCCCTTCTGGAGACCGCCAGAGCTGGCATGATGTTGGAG 540  
 QY 585 TTAGCCTTTCTGGCCATCGCCAGGAACCTGAAATACCGGGCCGGGATCAGGCGGATGAG 644  
 DB 541 TTAGCCTTTCTGGCCATCGCCAGGAACCTGAAATACCGGGCCGGGATCAGGCGGATGAG 600  
 QY 645 CCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAAAGCGCTCCAGCTGCTGCTCC 704  
 DB 601 CCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAAAGCGCTCCAGCTGCTGCTCC 660  
 QY 705 TTCATGTGAATCCAGGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGAGCGCTTCC 764  
 DB 661 TTCATGTGAATCCAGGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGAGCGCTTCC 720  
 QY 765 CCCTCCAGCGCTTCTTATTCAGAGGCTGAGCCAAATGGGAGGAAAGATGAGGAGCTC 824  
 DB 721 CCCTCCAGCGCTTCTTATTCAGAGGCTGAGCCAAATGGGAGGAAAGATGAGGAGCTC 780  
 QY 825 ACTGACACGCGCTTCTTACGAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTCGGT 884  
 DB 781 ACTGACACGCGCTTCTTACGAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTCGGT 840  
 QY 885 CTCCTCCGATCCACAGGAGGCTTAAACACTTACTTATTTATAGTACATATTTA 944  
 DB 841 CTCCTCCGATCCACAGGAGGCTTAAACACTTACTTATTTATAGTACATATTTA 900  
 QY 945 ATACCAAAAGGCGCTGATCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTG 1004  
 DB 901 ATACCAAAAGGCGCTGATCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTG 960  
 QY 1005 CTTTCTAGGACTTGGGGGCGCGCTCCCTCTCTAGCATACAAAGGTGGTGTGCTCC 1064  
 DB 961 CTTTCTAGGACTTGGGGGCGCGCTCCCTCTCTAGCATACAAAGGTGGTGTGCTCC 1020  
 QY 1065 AGCTACGCCCCAGGGGACACAGATGCATTTGGGGGTGAGGCGAGGTAACTGCTCCATCG 1124  
 DB 1021 AGCTACGCCCCAGGGGACACAGATGCATTTGGGGGTGAGGCGAGGTAACTGCTCCATCG 1080  
 QY 1125 CACCTCTAGTTTACGTTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCTCCGCTCT 1184  
 DB 1081 CACCTCTAGTTTACGTTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCTCCGCTCT 1140  
 QY 1185 CCAGGAGCTTATCTTCCGCCCATCTCCCAATAAGTGGGCGCTTGTCTGTAGGAGAC 1244  
 DB 1141 CCAGGAGCTTATCTTCCGCCCATCTCCCAATAAGTGGGCGCTTGTCTGTAGGAGAC 1200  
 QY 1245 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGAGGGGGAGGACAAAGGGCAGAGAG 1304  
 DB 1201 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGAGGGGGAGGACAAAGGGCAGAGAG 1260  
 QY 1305 TAGGCTCTAGCTGGCTTCTCTGGCCTTACTTAACACCCCCCTCGAG 1350  
 DB 1261 TAGGCTCTAGCTGGCTTCTCTGGCCTTACTTAACACCCCCCTCGAG 1306



## RESULT 5

AAH75184

ID AAH75184 standard; DNA; 1109 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Nucleotide sequence of human 32712 G-protein.

Human: G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder; congenital anomaly; pulmonary congestion; oedema; haemorrhage; adult respiratory distress syndrome; Goodpasture's syndrome; chronic obstructive pulmonary disease; asthma; pulmonary hypertension; liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis; cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure; brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage; acute meningitis; Parkinson's disease; Alzheimer's disease; glioma; chronic bacterial meningoencephalitis; multiple sclerosis; amyotrophic lateral sclerosis; stroke; Huntington's disease; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 124..699

/\*tag= a

/product= "G-protein"

WO200164887-A2.

07-SEP-2001.

27-FEB-2001; 2001WO-US06292.

29-FEB-2000; 2000US-0185606.

(MILL-) MILLENNIUM PHARM INC.

Meyers RA;

WPI; 2001-550182/61.

P-PSDB; AAG67156.

Novel human small G-protein polypeptides and polynucleotides for

treating lung disorders, liver disorders and brain disorders

Claim 2; Fig 26; 151pp; English.

The present sequence encodes a human G-protein. The specification

describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The

G-protein polypeptides and polynucleotides are useful as a target for

diagnosis and treatment of G-protein mediated or related disorders,

and for identifying agonists and antagonists for diagnosis and

treatment. They are useful for treating disorders of lung (e.g.

congenital anomalies, pulmonary congestion, oedema, adult respiratory

distress syndrome, haemorrhage, chronic obstructive pulmonary disease,

asthma, Goodpasture's syndrome and pulmonary hypertension), liver

(e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,

Wilson's disease, autoimmune hepatitis and hepatic failure), and

brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute

meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic

bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral

sclerosis, stroke and Huntington's disease).

Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

Query Match 40.4%; Score 1079; DB 22; Length 1109;

Best Local Similarity 99.5%; Pred. No. 2.2e-197;

Matches 1082; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

37 GGGACATGAGGGGACCCAGGCGCGTGGCCACCCCGGATGGGAGGCCCGGCGT 96

Db 23 GCGTCCGGAGGGGACGCCAGGCGCGTGGCCACCCCGGGATGGCGAGGCCCGCGAGCGCT 82  
QY 97 CCCC GCCCTGCGAGCTCCGAGCTACGAGCTACGCGGCAAGGTGATGCTTCTGGGAGACACAG 156  
Db 83 CCCC GCCCTGCGAGCTCCGAGCTACGAGCTACGCGGCAAGGTGATGCTTCTGGGAGACACAG 142  
QY 157 GCGTCCGGCAAAACATGTTTCTCTGATCCAAATTCAAAGACGGGGCTTCTCTCCGGAACCT 216  
Db 143 GCGTCCGGCAAAACATGTTTCTCTGATCCAAATTCAAAGACGGGGCTTCTCTCCGGAACCT 202  
QY 217 TCATAGCCACCGTCGCGATAGACTTCAGGAACAAGGTGGTGTGATGGATGGCGTGAGAG 276  
Db 203 TCATAGCCACCGTCGCGATAGACTTCAGGAACAAGGTGGTGTGATGGATGGCGTGAGAG 262  
QY 277 TGAAGCTGCAGATCTGGGACACCGCTGGGCGAGCAAGGTTCGCAAGCGGTCCACCATGCTT 336  
Db 263 TGAAGCTGCAGATCTGGGACACCGCTGGGCGAGCAAGGTTCGCAAGCGGTCCACCATGCTT 332  
QY 337 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCACCACAAATCTTCTTCG 396  
Db 323 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCACCACAAATCTTCTTCG 382  
QY 397 ACAACATCAGGCGCTGCTCAGTGCATGAGATTCATGATGATGCCAGAGGAGCTGGTATCA 456  
Db 383 ACAACATCAGGCGCTGCTCAGTGCATGAGATTCATGATGATGCCAGAGGAGCTGGTATCA 442  
QY 457 TGCTGCTAGGCAACAAGGGGATGATGAGCAGCAAGAGTGTATGCTTCCGAAGACGGAG 516  
Db 443 TGCTGCTAGGCAACAAGGGGATGATGAGCAGCAAGAGTGTATGCTTCCGAAGACGGAG 502  
QY 517 AGACCTTGCGCCAGGAGTAGCGGTGTTCCCTCTCTGGAGACCGGCAAGACTGGCATGA 576  
Db 503 AGACCTTGCGCCAGGAGTAGCGGTGTTCCCTCTCTGGAGACCGGCAAGACTGGCATGA 562  
QY 577 ATGTGGAGTTAGCTTCTTGCCATCGCCAAAGAACTGAAATACCGGCGCGGCATCAGG 636  
Db 563 ATGTGGAGTTAGCTTCTTGCCATCGCCAAAGAACTGAAATACCGGCGCGGCATCAGG 622  
QY 637 CGGATGAGCCCGAGCTTCCAGATCCGAGACTATATAGAGTCCCAAGAAAGCGTCCAGCT 696  
Db 623 CGGATGAGCCCGAGCTTCCAGATCCGAGACTATATAGAGTCCCAAGAAAGCGTCCAGCT 682  
QY 697 GCTGCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCAACAGGATGC 756  
Db 683 GCTGCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCAACAGGATGC 742  
QY 757 AGCTTCCCGCTCCAGGCGCTGCTTATTCCAAGAGGCTGAGCCAAATGGGAGAAAGATG 816  
Db 743 AGCTTCCCGCTCCAGGCGCTGCTTATTCCAAGAGGCTGAGCCAAATGGGAGAAAGATG 802  
QY 817 GAGGACTCACTGCACAGCCGCTTCTTAGCAGGAGGCTATATCTCCAACTCCTACTTGA 876  
Db 803 GAGGACTCACTGCACAGCCGCTTCTTAGCAGGAGGCTATATCTCCAACTCCTACTTGA 862  
QY 877 CTTGCGGTCTCCCGCATCCACAGGAGGCTAAACACTTAGCTTTATTTTAAATAGTAC 936  
Db 863 CTTGCGGTCTCCCGCATCCACAGGAGGCTAAACACTTAGCTTTATTTTAAATAGTAC 922  
QY 937 ATATTTTAAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 996  
Db 923 ATATTTTAAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 982  
QY 997 CTTTGTCTTCTTAGGACTTTGGGGGCGCGCTTCTCTCTTAAGCATATAAAGGTGGT 1056  
Db 983 CTTTGTCTTCTTAGGACTTTGGGGGCGCGCTTCTCTCTTAAGCATATAAAGGTGGT 1042  
QY 1057 GTTGTCTCAGCTCAGGCGGACACAGATGCATCTTGGGGGTGAGGCGAGTATGA 1116  
Db 1043 GTTGTCTCAGCTCAGGCGGACACAGATGCATCTTGGGGGTGAGGCGAGTATGA 1102  
QY 1117 CTCATC 1123  
|||||

Db 1103 CTCATC 1109

RESULT 6  
ABN83696  
ID ABN83696 standard; cDNA; 875 BP.  
XX  
AC ABN83696;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Human Ras protein 3 (RASP-3) cDNA.  
XX  
KW Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;  
KW immunosuppressive; antiinflammatory; signal transduction; gene;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 93..668  
FT /\*tag= a  
FT /product= "RASP-3"  
XX  
PN US6391580-B1.  
XX  
PD 21-MAY-2002.  
XX  
PF 08-MAY-1998; 98US-0075454.  
XX  
PR 12-DEC-1996; 96US-0766551.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Hillman JL, Tang YF, Lal P, Guegler KJ, Corley NC, Patterson C;  
PI Batra S, Baughn MR;  
XX  
DR WPI: 2002-498774/53.  
XX  
DR P-PSDB; ABB76426.  
XX  
PT New isolated polynucleotides encoding Ras proteins designated RASP-1  
PT and RASP-4, for diagnosing, preventing and treating disorders  
PT associated with cell proliferation, particularly cancer and immune  
PT disorders  
XX  
PS Example; Column 57-60; 34pp; English.  
XX  
CC The present sequence is that of a cDNA clone encoding novel human  
CC Ras protein 3 (RASP-3, see ABB76426). Nucleic acids encoding  
CC RASP-3 were initially identified in Incyte Clone 1528559 from a  
CC mononuclear cell cDNA library (UCMCL5T01) using a computer search  
CC for amino acid sequence alignments. The present consensus  
CC sequence was derived from overlapping and/or extended nucleic  
CC acids in Incyte clone 1528559 and shot-gun sequence SAEA03135,  
CC SABC10396 and SAEC10855. RASP-3 shows homology to rat Rab26,  
CC Northern analysis showed expression of RASP-3 in haematopoietic  
CC and immunological cDNA libraries, all associated with inflammation  
CC and the immune response. The invention provides 7 novel human Ras  
CC proteins (RASP-1 to -7) and polynucleotides, expression vectors,  
CC host cells, antibodies, agonists and antagonists. It also provides  
CC methods for diagnosing, treating or preventing disorders associated  
CC with RASP expression, especially cancer and immune disorders. A  
CC fragment of the present sequence, from about nucleotide 92 to about  
CC nucleotide 153, is useful as a hybridisation probe.  
XX  
SQ Sequence 875 BP; 203 A; 247 C; 260 G; 165 T; 0 other;  
Query Match 32.6%; Score 873; DB 24; Length 875;  
Best Local Similarity 100.0%; Pred. No. 5,7e-158;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
46 CGGCACGCCAGCGCGTGGCCACCGGATGGCGAGGCCCGCGCTCCCGCCCT 105  
|||||

Db 1 CGGCACGCCAGCGCGCTTCCACCGGGATGGCGAGGCCCGCGAGCGCTCCCGCCCT 60  
QY 106 GCAGTCCGAGCTAGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGGTGCGCA 165  
Db 61 GCAGTCCGAGCTAGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGGTGCGCA 120  
QY 166 AACATGTTTCTGATCCAAATTCAGGACGCGGCTTCTCTCGGGAACCTTCATAGCCA 225  
Db 121 AACATGTTTCTGATCCAAATTCAGGACGCGGCTTCTCTCGGGAACCTTCATAGCCA 180  
QY 226 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGAATGGGTGGAGTGAAGCTGC 285  
Db 181 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGAATGGGTGGAGTGAAGCTGC 240  
QY 286 AGATCTGGGACACCGCTGGGAGGAGGTTCCGAAGCGTCCACCATGCTTATTACAGAG 345  
Db 241 AGATCTGGGACACCGCTGGGAGGAGGTTCCGAAGCGTCCACCATGCTTATTACAGAG 300  
QY 346 ATGCTCAGGCGCTTCTTCTGCTGATGACATCACCAACAATCTTCTTCGACCAATCA 405  
Db 301 ATGCTCAGGCGCTTCTTCTGCTGATGACATCACCAACAATCTTCTTCGACCAATCA 360  
QY 406 GGGCTGGCTCAGTACGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 465  
Db 361 GGGCTGGCTCAGTACGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 466 GCAACAAGGCGGATATGAGCAGGAGGAGGATGATGATGATGATGATGATGATGATGATGAT 525  
Db 421 GCAACAAGGCGGATATGAGCAGGAGGAGGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 526 CCAGGAGTACGGTGTTCCTTCTGAGACGACGACGACGACGACGACGACGACGACGACGAC 585  
Db 481 CCAGGAGTACGGTGTTCCTTCTGAGACGACGACGACGACGACGACGACGACGACGACGAC 540  
QY 586 TAGCCTTCTGGCCATCGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 645  
Db 541 TAGCCTTCTGGCCATCGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 600  
QY 646 CCAGCTTCCAGATCCGAGACTATGATAGTCCCAAGAGAGCGCTCCAGCTGCTGCTCT 705  
Db 601 CCAGCTTCCAGATCCGAGACTATGATAGTCCCAAGAGAGCGCTCCAGCTGCTGCTCT 660  
QY 706 TCATGTGAATCCAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 765  
Db 661 TCATGTGAATCCAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
QY 766 CCTCCAGGCGCTGGCTTATTCAGAGGCTGAGGCAATGGGAGAGAGATGGAGGAGCTCA 825  
Db 721 CCTCCAGGCGCTGGCTTATTCAGAGGCTGAGGCAATGGGAGAGAGATGGAGGAGCTCA 780  
QY 826 CTGCACAGCGCTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885  
Db 781 CTGCACAGCGCTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
QY 886 TCCCGGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918  
Db 841 TCCCGGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 873

RESULT 7  
AAA52610  
ID AAA52610 standard; DNA; 797 BP.  
XX  
AC AAA52610;  
XX  
DT 07-DEC-2000 (first entry)  
XX  
DE Eosinophil activating peptide gene 8.  
XX  
KW Eosinophil activation; human; allergy; eosinophilia; cancer;  
KW inflammation; signalling peptide; cell adhesion peptide;  
KW G-protein coupled receptor; secreted protein; cell surface protein; ds.  
XX





```
QY 673 AGTCCAGAGAGAGCGCTCCAGCTGCTGCTCTTCTCATGTCAATCCAGGGGCGAGAGG 732
|||||
Db 659 AGTCCAGAGAGAGCGCTCCAGCTGCTGCTCTTCTCATGTCAATCCAGGGGCGAGAGG 718
|||||
QY 733 AGGCTCTGGAGGACAC-ACAGGATGACAGCCCTTCCCTCCAGGCC 776
|||
Db 719 AGGGCCANAGACCCCTGGGGATGCGAGTACTCCACTGCCACACC 763

RESULT 10
ABA02773
ID ABA02773 standard; DNA; 576 BP.
XX
AC ABA02773;
XX
DT 07-FEB-2002 (first entry)
XX
DE Mouse degranulation regulator encoding DNA SEQ ID NO 2.
XX
KW Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..576 a
FT /tag= a
FT /transl_except= (pos:247..249,aa:Lys)
FT /product= "degranulation regulator"
XX
PN WO200179478-A1.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-JP03268.
XX
PR 19-APR-2000; 2000JP-0118408.
XX
PA (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI Yamada T, Ido M;
XX
DR WPI; 2002-041335/05.
DR P-PSDB; AAM52189.
XX
PT Mast cell degranulation controller for treatment of allergies
PS Claim 10; Page 63-64; 85pp; Japanese.
XX
CC The invention relates to a protein for regulating degranulation of mast
CC cells (degranulation regulators) and the encoding polynucleotides, with
CC antiallergic activity, used in the treatment of allergies associated
CC with degranulation of mast cells.
XX
SQ Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Query Match 18.0%; Score 480; DB 24; Length 576;
Best Local Similarity 89.6%; Pred. No. 8.6e-83;
Matches 516; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 138 ATGCTTCTGGGAGACACAGCGCTCGCAAAACATGTTCTCATGCAATTCAAAGACGGG 197
|||||
Db 1 ATGCTTCTGGGAGATCGGGCGCTCGCAAAACCTGTTCTCATGCAATTCAAAGACGGG 60
|||||
QY 198 GCCTTCTGTCGGGAACCTTCATAGCCACCGCTCGCATAGACTTCAGGAACAAGTGGTG 257
|||||
Db 61 GCCTTCTGTCGGGAACCTTCATAGCCACCGCTCGCATAGACTTCAGGAATTAAGTGGTG 120
|||||
QY 258 ACTGTGGATGGCTGAGAGTGAAGCTGCAGATCTGGACACCGCTGGGAGGAACGGTTC 317
|||||
Db 121 ACAGTGGATGGTGCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTT 180
|||||
QY 318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTCTGCTGATGACATC 377
|||||
```

```
Db 181 CGAAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTGTCTCTGTGATGACATC 240
|||||
QY 378 ACCAACAATCTCTTTTCGACACATCAGGGCTGGCTCAGTCTGAGATTCATGAGTATGCC 437
|||||
Db 241 ACCAACCACTCTCTTTTCGACACATCAGGGCTGGCTCAGAGATTCATGAGTATGCC 300
|||||
QY 438 CAGAGGACGTGGTGTATCATGTCTGTAGGACACAAAGCGGATATAGCAGCGAAGAGTG 497
|||||
Db 301 CAGAGAGACGTGGTGTATCATGTCTGTAGGACACAAAGCGGATATAGCAGCGAAGAGTG 360
|||||
QY 498 ATCCCTTCGGAAGACGAGAGACCTTGGCCAGGGAGTACGGTCTTCCCTTCTGGAGACC 557
|||||
Db 361 ATCCGTTCTGAAGATGGAGACACATGGCCAGGGAATATGGTGTCTTCTCATGAGAGACC 420
|||||
QY 558 AGCGCCAGACTGGCATGATGTGGAGTTAGCTTTCTGCGCATCCCAAGGAACGAA 617
|||||
Db 421 AGTGCCAGACTGGCATGATGTGGAGTTAGCTTTCTGCGCATTCGCAATTCGCAAGGAACGAA 480
|||||
QY 618 TACCGGGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
|||||
Db 481 TACCGTGCAGGAGGAGCGCTGATGAGCCAGCTTCCAGATCCGAGACTATGTGGAGTCC 540
|||||
QY 678 CAGAGAAGCGCTCCAGCTGCTGCTCTTCATGCTGA 713
|||||
Db 541 CAGAGAAGCGCTCCAGCTGCTGCTCTTCCTTTGTGTGA 576
```

## RESULT 11

```
AAS81431
ID AAS81431 standard; cDNA; 475 BP.
XX
AC AAS81431;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #17235.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
DR P-PSDB; ABG17244.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID NO 17235; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
```

CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS94197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;

Query Match 17.5%; Score 468.6; DB 23; Length 475;  
 Best Local Similarity 99.2%; Pred. No. 1.3e-80;  
 Matches 471; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GGCACGTCTCACCTCTGTCAGGACATGACGGCAGCCAGCGCGGTGCCACCCGG 74  
 DB 1 GGAATTCCTCACCTCTGTCAGGACATGACGGCAGCCAGCGCGGTGCCACCCGG 60

QY 75 GATGGCGAGCCCGCCGAGCGCTCCCGCCCTGAGTCCGAGCTACGACCTCAGCGGCAAG 134  
 DB 61 GATGGCGAGCCCGCCGAGCGCTCCCGCCCTGAGTCCGAGCTACGACCTCAGCGGCAAG 120

QY 135 GTGATGCTTCTGGGACACACAGCGCTGGGCAAAACATGTTCTGATCCCAATTCAAAGAC 194  
 DB 121 GTGATGCTTCTGGGACACACAGCGCTGGGCAAAACATGTTCTGATCCCAATTCAAAGAC 180

QY 195 GGGGCTTCTGTCGGAACTTCATAGCACCGTCGGCATAGACTTCAGGAACAAGTG 254  
 DB 181 GGGGCTTCTGTCGGAACTTCATAGCACCGTCGGCATAGACTTCAGGAACAAGTG 240

QY 255 GTGACTGTGATGCGTGAGAGTGAAGCTGCATCTGGACACCGCTGGGAGGACCG 314  
 DB 241 GTGACTGTGATGCGTGAGAGTGAAGCTGCATCTGGACACCGCTGGGAGGACCG 300

QY 315 TTCGGAAGCTCACCCATGCTTATTACAGAGATGCTCAGGCTTGTCTGCTGTATGAC 374  
 DB 301 TTCGGAAGCTCACCCATGCTTATTACAGAGATGCTCAGGCTTGTCTGCTGTATGAC 360

QY 375 ATCACCACAAATCTCTTCGACACATCAGGCGCTGCTACTGAGATTCATGATAT 434  
 DB 361 ATCACCACAAATCTCTTCGACACATCAGGCGCTGCTACTGAGATTCATGATAT 420

QY 435 GCCCAGAGGACGTGGTGATGCTGCTAGGCAACAAAGCGGATATGACACCG 489  
 DB 421 GCCCAGAGGACGTGGTGATGCTGCTAGGCAACAAAGCGGATATGACACCG 475

## RESULT 12

AAV65197  
 ID AAV65197 standard; DNA; 1340 BP.

XX AC AAV65197;

XX 17-DEC-1998 (first entry)

DE Human RAB protein, SRAB, coding sequence.

XX Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;  
 KW Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;  
 KW viral infection; therapy; autoimmune sialosis; cystic fibrosis;  
 KW diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;  
 KW hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;  
 KW polycystic renal disease; ss.

OS Homo sapiens.

XX

Key Location/Qualifiers  
 CDS 43..615  
 FT /\*tag= a  
 FT /product= "SRAB"  
 XX WO9842839-A1.  
 XX 01-OCT-1998.  
 XX 25-MAR-1998; 98WO-US05871.  
 XX 26-MAR-1997; 97US-0824873.  
 XX (INCY-) INCYTE PHARM INC.  
 XX PA  
 XX Guegler KJ, Hillman JL;  
 XX WPI; 1998-542282/46.  
 XX P-PSDB; AAW80747.  
 PT New human RAB protein, SRAB - useful for treating disorders  
 PT associated with SRAB expression, including vesicle trafficking,  
 PT viral infection, and cancer  
 XX Claim 5; Fig 1; 31pp; English.  
 PS  
 CC This sequence encodes the human RAB protein, SRAB, of the invention. The  
 CC SRAB protein can be used in compositions useful in the diagnosis,  
 CC prevention, or treatment of disorders associated with vesicle trafficking  
 CC (including epilepsy, Huntington's disease, Parkinson's disease and  
 CC schizophrenia), cancer, or viral infection. Host cells containing the DNA  
 CC sequence can be used to produce SRAB recombinantly. The viral infections  
 CC that can be treated using SRAB include those caused by retroviruses  
 CC particularly HIV and HTLV, hepatitis viruses, particularly hepatitis C,  
 CC hantaviruses, herpesviruses, and arboviruses. The cancers that can be  
 CC treated include adenocarcinoma, leukaemia, lymphoma, melanoma and  
 CC sarcoma, particularly cancers of the endocrine, gastrointestinal and  
 CC nervous systems and cancers of the adrenal gland, brain, breast, colon,  
 CC oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,  
 CC prostate, salivary gland, stomach, thyroid, and uterus. The protein can  
 CC also be used to treat conditions such as autoimmune sialosis,  
 CC choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,  
 CC hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme  
 CC deficiency, pancreatitis, and polycystic renal disease.  
 XX  
 XX Sequence 1340 BP; 276 A; 387 C; 403 G; 274 T; 0 other;  
 Query Match 10.9%; Score 292.6; DB 19; Length 1340;  
 Best Local Similarity 70.9%; Pred. No. 6.8e-47;  
 Matches 416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY 116 CTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGCGCTGGCAAAACATGTTT 175  
 DB 21 CTACGACCTCGCCTTCAAGGTGATGCTGTTGGGGGACTCGGGTGGGAGACCTGTCT 80

QY 176 CCTGA-TCCAAATCAAGACGGGCGCTTCTGTCGGAGACCTTCATAGCACCGTCGGCA 234  
 DB 81 GCTGGGTGCGATTCAAGGATGGTGTCTTCTGGGGGGACCTTCATCTCCACCGT-AGCA 139

QY 235 TAGACTTCAGGAACAAGGTGGTGTGATGGCTGAGAGTGAAGCTGCAGATCTGGG 294  
 DB 140 TTGACTTCCGGACAAGTTCAGGCTGGATGGTGTGAAGTGAAGCTGCAGATCTGGG 199

QY 295 ACACCGCTGGGAGGAACCGTTCCGAGGCTCACCATGCTTATTACAGAGATGCTCAGG 354  
 DB 200 ACACAGCTGGTCAGGAGCGGTTCCCGAGTGTACCATGCTTACTACCGGGATGCTCATG 259

QY 355 CTTTGGTCTGCTGTATGATCATCACCAACAAATCTCTTTCGACACATCAGGCGCTGGC 414  
 DB 260 CTCTGCTGCTGCTCTACGATGTACCAACAAGCGCTCTCTTTCGACACATCAGGCGCTGGC 319

QY 415 TCACCTGAGATTCATGATGATGCCAGAGGACGTGGTGTATGCTGTGATGCTAGCAACAGG 474









|    |     |   |     |
|----|-----|---|-----|
| Db | 243 | GAGTATGGACTGCCCTTCATGGAGACCAGCGCCCAAGAGCGGGCCTCAAGTGGACTTGCC  | 302 |
| Qy | 591 | TTTTCTGGCCATCGCCCAAGGAACTGAATATACCGGGCCGGGATCAGCGGATGAGCCGAGC | 650 |
| Db | 303 | TTTCTGGCCATCGCCCAAGGAACTGAATATACCGGGCCGGGATCAGCGGATGAGCCGAGC  | 362 |
| Qy | 651 | TTCCAGATCCGAGACTATGTAGAGTCCAGAGAGGCGCTCCAGCTGCTGC             | 701 |
| Db | 363 | TTCCGGCTGCATGATTACGTTAAGAGGAGGTCGAGGGGCTCCTGCTGC              | 413 |

Search completed: June 23, 2003, 21:56:32  
 Job time : 570 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:32:44 ; Search time 3647 Seconds  
(without alignments)  
11874.618 Million cell updates/sec

Title: US-09-817-199b-1  
Perfect score: 2674  
Sequence: 1 ttcgcctgcggcgccgact.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

|            |               |   |  |       |               |   |  |
|------------|---------------|---|--|-------|---------------|---|--|
| Database : |               |   |  | EST:* |               |   |  |
| 1:         | em_estba:     | * |  | 1:    | em_estba:     | * |  |
| 2:         | em_esthum:    | * |  | 2:    | em_esthum:    | * |  |
| 3:         | em_estin:     | * |  | 3:    | em_estin:     | * |  |
| 4:         | em_estmu:     | * |  | 4:    | em_estmu:     | * |  |
| 5:         | em_estov:     | * |  | 5:    | em_estov:     | * |  |
| 6:         | em_estpl:     | * |  | 6:    | em_estpl:     | * |  |
| 7:         | em_estro:     | * |  | 7:    | em_estro:     | * |  |
| 8:         | em_htc:       | * |  | 8:    | em_htc:       | * |  |
| 9:         | gb_estl:      | * |  | 9:    | gb_estl:      | * |  |
| 10:        | gb_est2:      | * |  | 10:   | gb_est2:      | * |  |
| 11:        | gb_htc:       | * |  | 11:   | gb_htc:       | * |  |
| 12:        | gb_est3:      | * |  | 12:   | gb_est3:      | * |  |
| 13:        | gb_est4:      | * |  | 13:   | gb_est4:      | * |  |
| 14:        | gb_est5:      | * |  | 14:   | gb_est5:      | * |  |
| 15:        | em_estfun:    | * |  | 15:   | em_estfun:    | * |  |
| 16:        | em_estom:     | * |  | 16:   | em_estom:     | * |  |
| 17:        | gb_gss:       | * |  | 17:   | gb_gss:       | * |  |
| 18:        | em_gss_hum:   | * |  | 18:   | em_gss_hum:   | * |  |
| 19:        | em_gss_inv:   | * |  | 19:   | em_gss_inv:   | * |  |
| 20:        | em_gss_pln:   | * |  | 20:   | em_gss_pln:   | * |  |
| 21:        | em_gss_vrt:   | * |  | 21:   | em_gss_vrt:   | * |  |
| 22:        | em_gss_fun:   | * |  | 22:   | em_gss_fun:   | * |  |
| 23:        | em_gss_mam:   | * |  | 23:   | em_gss_mam:   | * |  |
| 24:        | em_gss_mus:   | * |  | 24:   | em_gss_mus:   | * |  |
| 25:        | em_gss_other: | * |  | 25:   | em_gss_other: | * |  |
| 26:        | em_gss_pro:   | * |  | 26:   | em_gss_pro:   | * |  |
| 27:        | em_gss_rod:   | * |  | 27:   | em_gss_rod:   | * |  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |             |              |             |  |  |  |
|------------|-------|-------------|--------------|-------------|--|--|--|
| Result No. | Score | Query Match | Length DB ID | Description |  |  |  |
| 1          | 921.6 | 34.5        | 1014 14      | BM921365    |  |  |  |
| 2          | 884.2 | 33.1        | 887 9        | AL522282    |  |  |  |
| C 3        | 847.4 | 31.7        | 855 9        | AL580781    |  |  |  |
| C 4        | 802.6 | 30.0        | 833 9        | AL522281    |  |  |  |
| 5          | 755.4 | 28.2        | 881 12       | BG759655    |  |  |  |
| 6          | 741.6 | 27.7        | 868 14       | BQ690583    |  |  |  |

|      |       |      |      |    |          |             |           |
|------|-------|------|------|----|----------|-------------|-----------|
| 7    | 733.2 | 27.4 | 761  | 13 | BI819064 | BI819064    | 603033417 |
| C 8  | 701.2 | 26.3 | 702  | 12 | BM981939 | BM981939    | UI-CF-EN1 |
| 9    | 701.2 | 26.2 | 1051 | 14 | BG283602 | BG283602    | 602407930 |
| C 10 | 696.8 | 26.1 | 700  | 14 | BQ028124 | UI-H-COO-   | BQ028124  |
| C 11 | 643.6 | 24.1 | 1069 | 12 | BG282782 | BG282782    | 602405867 |
| C 12 | 626.4 | 23.4 | 629  | 14 | BQ183276 | UI-H-EUO-   | BQ183276  |
| C 13 | 605.4 | 22.6 | 607  | 9  | AF188522 | AF188522    | AF188522  |
| 14   | 601.6 | 22.5 | 740  | 13 | BI767046 | BI767046    | 603054267 |
| 15   | 600.6 | 22.5 | 691  | 12 | BG253976 | BG253976    | 602366910 |
| 16   | 585.6 | 21.9 | 749  | 9  | AL559085 | AL559085    | AL559085  |
| 17   | 581.2 | 21.7 | 886  | 14 | BQ686932 | BQ686932    | AGENCOURT |
| 18   | 553.2 | 20.7 | 676  | 10 | BB598938 | BB598938    | BB598938  |
| 19   | 539.2 | 20.2 | 603  | 10 | AW955318 | AW955318    | EST367388 |
| 20   | 529   | 19.8 | 529  | 14 | BM706070 | UI-E-DWO-   | BM706070  |
| 21   | 527.6 | 19.7 | 651  | 10 | BB633978 | BB633978    | BB633978  |
| 22   | 516.8 | 19.3 | 655  | 10 | BB635649 | BB635649    | BB635649  |
| 23   | 514.6 | 19.2 | 522  | 13 | BM151643 | TCBAP1E10   | BM151643  |
| C 24 | 503.4 | 18.8 | 507  | 13 | BM151381 | TCBAP1D86   | BM151381  |
| 25   | 473.6 | 17.7 | 480  | 10 | AW270880 | x605e04.x   | AW270880  |
| 26   | 464.4 | 17.4 | 493  | 9  | AA324998 | AA324998    | EST28027  |
| 27   | 460.6 | 17.2 | 488  | 13 | BM149118 | BM149118    | TCAAP2E63 |
| 28   | 442.4 | 16.5 | 676  | 10 | BB206788 | BB206788    | BB206788  |
| 29   | 440.4 | 16.5 | 766  | 14 | BQ188402 | UI-E-EJ1-   | BQ188402  |
| 30   | 423.8 | 15.8 | 553  | 12 | BE755280 | 205068 MA   | BE755280  |
| 31   | 422.4 | 15.8 | 733  | 10 | BB619625 | BB619625    | BB619625  |
| 32   | 415.2 | 15.5 | 437  | 14 | W36293   | HFBEET-70 H | W36293    |
| 33   | 412.4 | 15.4 | 793  | 12 | BG256040 | BG256040    | 602367337 |
| C 34 | 409.4 | 15.3 | 434  | 9  | AA806096 | AA806096    | oc25b06.s |
| C 35 | 399.4 | 14.9 | 401  | 10 | AW768344 | AW768344    | hk73a07.x |
| 36   | 397.4 | 14.9 | 657  | 10 | BB196489 | BB196489    | BB196489  |
| C 37 | 385   | 14.4 | 401  | 9  | AA631787 | AA631787    | np68e10.s |
| C 38 | 381   | 14.2 | 403  | 9  | AA878511 | AA878511    | oj19b02.s |
| 39   | 380.4 | 14.2 | 455  | 10 | AA464449 | AA464449    | BP230015B |
| 40   | 377.8 | 14.1 | 385  | 9  | AI243836 | qh84d08.x   | AI243836  |
| 41   | 374.2 | 14.0 | 380  | 10 | BE241772 | BE241772    | TCAAP2E02 |
| 42   | 362.8 | 13.6 | 579  | 10 | BB621927 | BB621927    | BB621927  |
| 43   | 358   | 13.4 | 368  | 13 | BM152052 | BM152052    | TCBAP1E12 |
| 44   | 347   | 13.0 | 348  | 9  | AA310682 | AA310682    | EST181496 |
| 45   | 345.6 | 12.9 | 404  | 13 | BM149006 | BM149006    | TCAAP2E57 |

ALIGNMENTS

RESULT 1  
BM921365

LOCUS  
DEFINITION  
AGENCOURT\_6626159 NIH\_MGC\_115 Homo sapiens CDNA clone IMAGE:572779

5', mRNA sequence.

ACCESSION  
BM921365

VERSION  
BM921365.1

KEYWORDS  
EST.

SOURCE  
human.

ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 1014)

AUTHORS  
NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
Unpublished (1999)

COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13787 row: 1 column: 04  
High quality sequence stop: 735.  
Location/Qualifiers  
1. .1014

| BASE COUNT            | 237   | a  | 288          | c          | 295 | g      | 194   | t    | ORIGIN |
|-----------------------|-------|--|--------------|------------|-----|--------|-------|------|--------|
| Query Match           | 34.5% | Score  | 921.6;       | DB         | 14; | Length | 1014; |      |        |
| Best Local Similarity | 96.7% | Pred.  | No. 5.2e-93; |            |     |        |       |      |        |
| Matches               | 973;  | Conservative   | 0;           | Mismatches | 29; | Indels | 4;    | Gaps | 3;     |
| QY                    | 2     | TCGCTCGCGCGCGGCACTGCTCACCTCTCGTCCAGGACATCACGGCAGCCAGCCAGCGCGC  | 61           |            |     |        |       |      |        |
| DB                    | 8     | TCGCTCGCGCGCGGCACTGCTCACCTCTCGTCCAGGACATCACGGCAGCCAGCCAGCGCGC  | 67           |            |     |        |       |      |        |
| QY                    | 62    | CGTTGCCACCCGGGATGGGAGGCCCGGAGCGCTCCCGCCCTGCAGTCCGAGCTACGA      | 121          |            |     |        |       |      |        |
| DB                    | 68    | CGTTGCCACCCGGGATGGGAGGCCCGGAGCGCTCCCGCCCTGCAGTCCGAGCTACGA      | 127          |            |     |        |       |      |        |
| QY                    | 122   | CCTCAGCGGCAAGTGTATGCTTCTGGAGACACAGCGCTCGCAAAACATGTTTCTGTAT     | 181          |            |     |        |       |      |        |
| DB                    | 128   | CCTCAGCGGCAAGTGTATGCTTCTGGAGACACAGCGCTCGCAAAACATGTTTCTGTAT     | 187          |            |     |        |       |      |        |
| QY                    | 182   | CCAATTCAAAGACGGGCGCTTCTGTCCGAACTTCATAGCCACCGTCGGCATAGACTT      | 241          |            |     |        |       |      |        |
| DB                    | 188   | CCAATTCAAAGACGGGCGCTTCTGTCCGAACTTCATAGCCACCGTCGGCATAGACTT      | 247          |            |     |        |       |      |        |
| QY                    | 242   | CAGGAACAAGGTGTGACTGTGGATGGCGCTGAGAGTGAAGCTGCAGATCTGGACACCGC    | 301          |            |     |        |       |      |        |
| DB                    | 248   | CAGGAACAAGGTGTGACTGTGGATGGCGCTGAGAGTGAAGCTGCAGATCTGGACACCGC    | 307          |            |     |        |       |      |        |
| QY                    | 302   | TGGGCAGGAACGGTTCGGAAGCGGTCAACCCATGCTTATTACAGAGATGCTCAGGCGCTTGT | 361          |            |     |        |       |      |        |
| DB                    | 308   | TGGGCAGGAACGGTTCGGAAGCGGTCAACCCATGCTTATTACAGAGATGCTCAGGCGCTTGT | 367          |            |     |        |       |      |        |
| QY                    | 362   | TCTGCTGTATGACATACCAACAATCTTTTCGACAAACATCAGGGCGTGGCTCACTGA      | 421          |            |     |        |       |      |        |
| DB                    | 368   | TCTGCTGTATGACATACCAACAATCTTTTCGACAAACATCAGGGCGTGGCTCACTGA      | 427          |            |     |        |       |      |        |
| QY                    | 422   | GATTTCATGATGTGCCAGAGGAGCTGTGTATCATGCTGTAGGCACACAAAGCGCGATAT    | 481          |            |     |        |       |      |        |
| DB                    | 428   | GATTTCATGATGTGCCAGAGGAGCTGTGTATCATGCTGTAGGCACACAAAGCGCGATAT    | 487          |            |     |        |       |      |        |
| QY                    | 482   | GAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCACGGGAGTACGGTGT  | 541          |            |     |        |       |      |        |
| DB                    | 488   | GAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCACGGGAGTACGGTGT  | 547          |            |     |        |       |      |        |
| QY                    | 542   | TCCCTTCTGGAGACAGGGCCAGACGTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT     | 601          |            |     |        |       |      |        |
| DB                    | 548   | TCCCTTCTGGAGACAGGGCCAGACGTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT     | 607          |            |     |        |       |      |        |
| QY                    | 602   | CGCCAAGGAACTGAAATACCGGGCGGGCATCAGCGGATGAGCCCACTTCCAGATCCG      | 661          |            |     |        |       |      |        |
| DB                    | 608   | CGCCAAGGAACTGAAATACCGGGCGGGCATCAGCGGATGAGCCCACTTCCAGATCCG      | 667          |            |     |        |       |      |        |
| QY                    | 662   | AGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGTCTCTTCATGTGAATCCCAAG       | 721          |            |     |        |       |      |        |
| DB                    | 668   | AGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGTCTCTTCATGTGAATCCCAAG       | 727          |            |     |        |       |      |        |
| QY                    | 722   | GGGCAGAGAGGCGCTCTGGAGGCACACAGATGAGCCCTTCCCGCTCCCAAGCGCTGGCT    | 781          |            |     |        |       |      |        |
| DB                    | 728   | GGGCAGAGAGGCGCTCTGGAGGCACACAGATGAGCCCTTCCCGCTCCCAAGCGCTGGCT    | 787          |            |     |        |       |      |        |

```
Db 121 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCCAATTCAAAGAC 180
Qy 195 GGGGCTTCTGTCGCGAACCTTTCATAGCCACCGTGGCATAGACTTTCAGGAAACAGGTG 254
Db 181 GGGGCTTCTGTCGCGAACCTTTCATAGCCACCGTGGCATAGACTTTCAGGAAACAGGTG 240
Qy 255 GTGACTGTGATGCGGTGAGAGTGAAGCTGAGATCGGACATCGGACACCGTGGCGAGGACGG 314
Db 241 GTGACTGTGATGCGGTGAGAGTGAAGCTGAGATCGGACATCGGACACCGTGGCGAGGACGG 300
Qy 315 TTCCGAAGCTTCACCCATGCTTATACAGAGATGCTCAGGCTTTCGCTGCTGTATGAC 374
Db 301 TTCCGAAGCTTCACCCATGCTTATACAGAGATGCTCAGGCTTTCGCTGCTGTATGAC 360
Qy 375 ATCACCACAAATCTTCTTCGACAAACATCAGGCGCTTTCGCTGCTGTATGAGTAT 434
Db 361 ATCACCACAAATCTTCTTCGACAAACATCAGGCGCTTTCGCTGCTGTATGAGTAT 420
Qy 435 GCCCAGAGGACGTGTGATCATGCTGCTAGGCAACAGCGGATATGACAGGAAAGA 494
Db 421 GCCCAGAGGACGTGTGATCATGCTGCTAGGCAACAGCGGATATGACAGGAAAGA 480
Qy 495 GTGATCGCTTCCGAAGACGAGAGACCTTGGCCAGGAGTACGGTGTTCCTCTCGAG 554
Db 481 GTGATCGCTTCCGAAGACGAGAGACCTTGGCCAGGAGTACGGTGTTCCTCTCGAG 540
Qy 555 ACCAGCGCCCAAGACTGGCATGAATGTGGAGTATAGCTTTCCTGCGCATCGCAAGAACTG 614
Db 541 ACCAGCGCCCAAGACTGGCATGAATGTGGAGTATAGCTTTCCTGCGCATCGCAAGAACTG 600
Qy 615 AATACCGCGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674
Db 601 AATACCGCGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 660
Qy 675 TCCAGAAAGAGCGCTCAGCTGCTCTTCATGTGAATCCAGGCGGAGAGAGAG 734
Db 661 TCCAGAAAGAGCGCTCAGCTGCTCTTCATGTGAATCCAGGCGGAGAGAGAG 720
Qy 735 GCTCTGGAGGACACAGAGATGAGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 794
Db 721 GCTCTGGAGGACACAGAGATGAGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 780
Qy 795 TGAGCCATGGGAGAGAGATGAGGAGTCACTGACAGCGCTTCCAGAGGAGCTA 854
Db 781 TGAGCCATGGGAGAGAGATGAGGAGTCACTGACAGCGCTTCCAGAGGAGCTA 840
Qy 855 TACTCCAACTCTACTTGTAGTTCCTGCGGCTCCTCCCGCATCCACAGG 901
Db 841 TACTCCAACTCTACTTGTAGTTCCTGCGGCTCCTCCCGCATCCACAGG 887
```

```
RESULT 3
AL580781/c
LOCUS
DEFINITION AL580781 LTI_NFL008.TC2 Homo sapiens cDNA clone CS0DJ014YB11 3
prime, mRNA sequence.
ACCESSION AL580781
VERSION AL580781.1 GI:12947137
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 855)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
```

```
source 1. .855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DJ014YB11"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/notes="vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 204 a 224 c 224 g 201 t 2 others
ORIGIN
Query Match 31.7%; Score 847.4; DB 9; Length 855;
Best Local Similarity 99.6%; Pred. No. 8.5e-85;
Matches 848; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1700 ATTTTGTAGACACACAGGTTTCCCATGTTGCCAGGCTGCTTGAATTCCTGAGCTC 1759
Db 855 ATTTTGTAGACACAGGTTTCCCATGTTGCCAGGCTGCTTGAATTCCTGAGCTC 796
Qy 1760 AAGCAACTCGCGGCTCGGCTCCCAAGTACTGGATTACACGAGAGGACCAATGC 1819
Db 795 AAGCAACTCGCGGCTCGGCTCCCAAGTACTGGATTACACGAGAGGACCAATGC 736
Qy 1820 CCAGGTAGATGTCTTATCCCAATCTTTGGCAGGATGCAGCTCCACAGGCGATTC 1879
Db 735 CCAGGTAGATGTCTTATCCCAATCTTTGGCAGGATGCAGCTCCACAGGCGATTC 676
Qy 1880 TTCAAGCAGCTGAAGTGTTAGCCCTCTGGGTTAAGAGCCAGATAGAGGAAATCCCTT 1939
Db 675 TTCAAGCAGCTGAAGTGTTAGCCCTCTGGGTTAAGAGCCAGATAGAGGAAATCCCTT 616
Qy 1940 TCCTAGTGTGGAATGTGTGAAAAAAGAGAAATCCCTGCTCTGAGCTGGTGG 1999
Db 615 TCCTAGTGTGGAATGTGTGAAAAAAGAGAAATCCCTGCTCTGAGCTGGTGG 556
Qy 2000 GAGACAGATTAAGCAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTC 2059
Db 555 GAGACAGATTAAGCAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTC 496
Qy 2060 TCCCTGACCCCATGCCCTTTCCTTTAACTTCTCAACACATACCAGGCGCTAACTGC 2119
Db 495 TCCCTGACCCCATGCCCTTTCCTTTAACTTCTCAACACATACCAGGCGCTAACTGC 436
Qy 2120 TTTACCTCCCTCTACTGAGTCAAGTTAGTGGGAGGTCACCCATTTCCGAGTTAA 2179
Db 435 TTTACCTCCCTCTACTGAGTCAAGTTAGTGGGAGGTCACCCATTTCCGAGTTAA 376
Qy 2180 ACCAATGCATATCAGTAAACAAAGTCATGTGGGTATGTCTGGGTAGAGAGGGGTA 2239
Db 375 ACCAATGCATATCAGTAAACAAAGTCATGTGGGTATGTCTGGGTAGAGAGGGGTA 316
Qy 2240 GCAAGTTTCATGTCTCTCTTGTGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGA 2299
Db 315 GCAAGTTTCATGTCTCTCTTGTGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGA 256
Qy 2300 AGTGACAGGAAACATCAGGTCATGACCTGAGGAGCTCTTTTACTGCAGCTCTGCCGGCT 2359
Db 255 AGTGACAGGAAACATCAGGTCATGACCTGAGGAGCTCTTTTACTGCAGCTCTGCCGGCT 196
Qy 2360 GGAGGGGAGAGGGGGAGGAGAAAGTATGCGCTGCACATTTCTGAGGCTACTGCAATTC 2419
Db 195 GGAGGGGAGAGGGGGAGGAGAAAGTATGCGCTGCACATTTCTGAGGCTACTGCAATTC 136
Qy 2420 TTTCAAGGCAGAAATCTTGTCTGAGCAGTCCAGGCTCCAGTTTGGCCCGCAAGAA 2479
```

```

135 TTTCAAGCCAGAAATCTGCTGAGCAGTACAGGCTCCAGTTTGGGCCCATTAAGGA 76
2480 GTTCTCCGTGGCTCCCTCAGCAGAGGAGGAGGCTGACATTTGCCAGTCTCTTG 2539
75 GTTCTCCGTGGCTCCCTCAGCAGAGGAGGAGGAGGCTGACATTTGCCAGTCTCTTG 16
2540 GGGCCCAAGGC 2550
15 GGGCCCAAGGC 5

RESULT 4
AL522281/c
LOCUS
DEFINITION
AL522281 LTI_NFL004_NBC2 833 bp mRNA linear EST 13-FEB-2001
prime, mRNA sequence.
ACCESSION
AL522281
VERSION
AL522281.1 GI:12785774
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB008YF08"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 206 a 207 c 215 g 200 t 5 others
ORIGIN

Query Match 30.0%; Score 802.6; DB 9; Length 833;
Best Local Similarity 98.7%; Pred. No. 7.3e-80;
Matches 814; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1673 GTACATCACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTCGCCATGTTC 1732
:|||||
DB 825 KTACCATCACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTCGCCATGTTC 766
:|||||
QY 1733 CCAGGCTGCTTGAATTCCTGAGCTCAAGCAACTCGCGGCTCGGCTCCCAAGTAC 1792
:|||||
DB 765 CCAGGCTGCTTGAATTCCTGAGCTCAAGCAACTCGCGGCTCGGCTCCCAAGTAC 706
:|||||
QY 1793 TGGATTACACGAGAGGACCAATGCCAGGCTAGATGTCTTATCCCAATCCTTTGG 1852
:|||||
DB 705 TGGATTACACGAGAGGACCAATGCCAGGCTAGATGTCTTATCCCAATCCTTTGG 646
:|||||
QY 1853 CAGCATCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTAGCCCTCCTGGT 1912
:|||||
DB 645 CAGCATCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTAGCCCTCCTGGT 586
:|||||

```

```

QY 1913 TAAGAGCCAGATAAGGAGAAATCCCTTCTCCTAGGTTTGAATGTGTGTGAAAAAAGA 1972
|||||
DB 585 TAAGAGCCAGATAAGGAGAAATCCCTTCTCCTAGGTTTGAATGTGTGTGAAAAAAGA 526
|||||
QY 1973 GAAATCCCTGGCTCCCTGGAGCTGGTGGAGACAAGATTAAAGCAAACTCCCTGACATGT 2032
|||||
DB 525 GAAATCCCTGGCTCCCTGGAGCTGGTGGAGACAAGATTAAAGCAAACTCCCTGACATGT 466
|||||
QY 2033 ATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACACCATGCGCTTTCCCTTTAACTTC 2092
|||||
DB 465 ATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACACCATGCGCTTTCCCTTTAACTTC 406
|||||
QY 2093 TCAACAGATACACAGGCTAAACTGCTTTACTCCCTCCCTCTACTG-AGTCAGGTTAGT 2151
|||||
DB 405 TCAACAGATACACAGGCTAAACTGCTTTACTCCCTCCCTGACACTGCACTCAGGTTAGT 346
|||||
QY 2152 GGTGGGAGTCAACCATTTCCGAGTTAAACCAATGATATAGTAAACAAGTCATGT 2211
|||||
DB 345 GGTGGGAGTCAACCATTTCCGAGTTAAACCAATGATATAGTAAACAAGTCATGT 286
|||||
QY 2212 GGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCCTCTTGGTCACATATC 2271
|||||
DB 285 GGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCCTCTTGGTCACATATC 226
|||||
QY 2272 TCCCAAAGCTCTGATCCCTGCCATGGAGTGGACAGGAACATGAGTCTAGCTGCA 2331
|||||
DB 225 TCCCAAAGCTCTGATCCCTGCCATGGAGTGGACAGGAACATGAGTCTAGCTGCA 166
|||||
QY 2332 GGCATCTTTACTGCAAGCTCTCGCGCTCGGGGGAGGGGAGGAGAGTATGCGC 2391
|||||
DB 165 GGCATCTTTACTGCAAGTCTCGCGCTCGGGGGAGGGGAGGAGAGTATGCGC 106
|||||
QY 2392 TGCACATTTCTGAGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTCTCTGAGCAGTCA 2451
|||||
DB 105 TGCACATTTCTGAGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTCTCTGAGCAGTCA 46
|||||
QY 2452 GCGGCTCAGTTTGGGCCCGGATAGGAAGTCTCCTCGTGGCTGCC 2496
|||||
DB 45 GCGGCTCAGTTTGGGCCCGGATAGGAAGTCTCCTCGTGGCTGCC 1

RESULT 5
BG759655 881 bp mRNA linear EST 15-MAY-2001
LOCUS 602713348F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853626 5',
DEFINITION mRNA sequence.
ACCESSION BG759655
VERSION BG759655.1 GI:14070308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1699 row: p column: 11
High quality sequence stop: 707.
Location/Qualifiers
1..881
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```



```
/clone="IMAGE:4853626"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: B-cells; Vector: pOB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      220 a 236 c 261 g 164 t
ORIGIN
Query Match      28.2%; Score 755.4; DB 12; Length 881;
Best Local Similarity 94.6%; Pred. No. 1.1e-74;
Matches 805; Conservative 0; Mismatches 41; Indels 5; Gaps 2;
Qy 566 GACTGGCATGAATGTGGAGTTAGCTTTCTGGCCATCCCAAGGAAGTGAATACCGGGC 625
Db 2 GACTGGCATGAATGTGGAGTTAGCTTTCTGGCCATCCCAAGGAAGTGAATACCGGGC 61
Qy 626 CGGCATCAGCGCATGAGCCGACCTCCAGATCCGAGACTATGTAGTCCCAAGAA 685
Db 62 CGGCATCAGCGCATGAGCCGACCTCCAGATCCGAGACTATGTAGTCCCAAGAA 121
Qy 686 CGCTCCAGCTGCTGCTTCCTTCATGTGAATCCAGGGGAGAGAGGCTCTGGAGGC 745
Db 122 CGCTCCAGCTGCTGCTTCCTTCATGTGAATCCAGGGGAGAGAGGCTCTGGAGGC 181
Qy 746 ACACAGGATGAGCGCTTCCCGCTCCAGGCGCTTATTCACAGAGGCTGAGCCAATGG 805
Db 182 ACACAGGATGAGCGCTTCCCGCTCCAGGCGCTTATTCACAGAGGCTGAGCCAATGG 241
Qy 806 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCCTAGCAGGAGGCTATCTCCACTC 865
Db 242 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCCTAGCAGGAGGCTATCTCCACTC 301
Qy 866 CTACTTGAGTCTCTCGGCTCTCCCGCATCCACAGGAGGCTAAACACTTAGCTTTAT 925
Db 302 CTACTTGAGTCTCTCGGCTCTCCCGCATCCACAGGAGGCTAAACACTTAGCTTTAT 361
Qy 926 TTTAATAGTACATAATTTAATACCAAAAGAGCGCCCTGGATCCCCAAAACCGAGGCTG 985
Db 362 TTTAATAGTACATAATTTAATACCAAAAGAGCGCCCTGGATCCCCAAAACCGAGGCTG 421
Qy 986 GGAGCTAGTGGCCCTTTTGTCTTAGSACTTGGGGGCGCGCCCTCCCTCTAAGCAT 1045
Db 422 GGAGCTAGTGGCCCTTTTGTCTTAGSACTTGGGGGCGCGCCCTCCCTCTAAGCAT 481
Qy 1046 ACAAGGTGTGTGTCTCCAGCTCAGCGCCAGGGGACACAGATGCACCTTTGGGGGTGAG 1105
Db 482 ACAAGGTGTGTGTCTCCAGCTCAGCGCCAGGGGACACAGATGCACCTTTGGGGGTGAG 541
Qy 1106 GCAGGTAACTGACTCCATCGACCTTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCGAGC 1165
Db 542 GCAGGTAACTGACTCCATCGACCTTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCGAGC 601
Qy 1166 CTTTCACGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCCATCTCCCAATAGTGGGCC 1225
Db 602 CTTTCACGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCCATCTCCCAATAGTGGGCC 661
Qy 1226 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGAGATATGAGATGGAGGGG 1285
Db 662 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGAGATATGAGAGGGGAGGG 721
Qy 1286 GAGGACAAAGGGGAGAGAGTGGTCTA --- GCTGGGTATCTCTGGGCTTACTTAACACC 1342
Db 722 GAGGACAAAGGGGAGAGAGTAAAGGACAAAGCATGGCAAGTCTCTGGCCGTACAAAGAGCG 781
Qy 1343 CCCTGG -- AGGCATGCCCCCTTTTCTCCAGGCACACAGCATATTGGGGCACCTTGAATAT 1400

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
782 CCCGGGAGGCGATCCCTTTATAAGCCAGGACACAGCCATTTGGGACCTGGAATAG 841
Qy 1401 TGGTTCCAGGC 1411
Db 842 TGGACCGGGAC 852

RESULT 6
BQ690583
LOCUS      BQ690583
DEFINITION AGENCOURT_8186311 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251642
5', mRNA sequence.
ACCESSION BQ690583
VERSION    BQ690583.1 GI:21815899
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 868)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2396 row: k column: 03
High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source           1..868
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6251642"
                     /clone_lib="NIH_MGC_110"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      215 a 239 c 240 g 172 t 2 others
ORIGIN
Query Match      27.7%; Score 741.6; DB 14; Length 868;
Best Local Similarity 93.6%; Pred. No. 3.5e-73;
Matches 805; Conservative 0; Mismatches 51; Indels 4; Gaps 3;
Qy 490 AAAGAGTGTATCCGTTCCGACGACGAGAGACTTGGCCAGGAGTACGCTTCCCTCC 549
Db 1 AAAGAGTGTATCCGTTCCGACGACGAGAGACTTGGCCAGGAGTACGCTTCCCTCC 60
Qy 550 TGGAGACAGCGCCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAGG 609
Db 61 TGGAGACAGCGCCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAGG 120
Qy 610 AACTGAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCAGATCCGAGACTATG 669
Db 121 AACTGAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCAGATCCGAGACTATG 180
Qy 670 TAGAGTCCCAAGAACGCTCCAGCTGCTCCTCTCATGTGAATCCAGGGGCGAGAG 729
Db 181 TAGAGTCCCAAGAACGCTCCAGCTGCTCCTCTCATGTGAATCCAGGGGCGAGAG 240
```

|            |   |  |                        |
|------------|---|--|------------------------|
| Qy         | 730   | AGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCCTCCAGGCTGGCTTATTC         | 789                    |
| Db         | 241   | AGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCCTCCAGGCTGGCTTATTC         | 300                    |
| Qy         | 790   | GAGGCTGAGCCATGGGAGGAAGATGGAGGACTCACTGCACAGCCGCTTCTAGCAGG     | 849                    |
| Db         | 301   | GAGGCTGAGCCATGGGAGGAAGATGGAGGACTCACTGCACAGCCGCTTCTAGCAGG     | 360                    |
| Qy         | 850   | AGCTATATCTCCAACTCCTACTTGAGTTCCTGGGTCTCCCGATCCACAGGAGGTAA     | 909                    |
| Db         | 361   | AGCTATATCTCCAACTCCTACTTGAGTTCCTGGGTCTCCCGATCCACAGGAGGTAA     | 420                    |
| Qy         | 910   | AACACTTAGCTTTATTTTAAATAGTACATAATTAATACCAAAAGGCGCTGGATCCC     | 969                    |
| Db         | 421   | AACACTTAGCTTTATTTTAAATAGTACATAATTAATACCAAAAGGCGCTGGATCCC     | 480                    |
| Qy         | 970   | CAAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCTTCTTAGGACTT - GGCGGGCGGC  | 1028                   |
| Db         | 481   | CAAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCTTCTTAGGACTT - GGCGGGCGGC  | 540                    |
| Qy         | 1029  | CCTCCCTCCTAAGCATCAACAAGGTGGTGTGTCTCCAGCTCAGCCCCAGGGGACACAGAT | 1088                   |
| Db         | 541   | CCTCCCTCCTAAGCATCAACAAGGTGGTGTGTCTCCAGCTCAGCCCCAGGGGACACAGAT | 600                    |
| Qy         | 1089  | GCATTTT - GGGGGTGGAGGAGGTAATGACTCCATCCACCTCAGTTCAGCTGGACAGA  | 1147                   |
| Db         | 601   | GCATTTTGGGGGGTGGAGGAGGTAATGACTCCATCCACCTCAGTTCAGCTGGACAGA    | 660                    |
| Qy         | 1148  | GGCTCAGGTGACCCAGCCCTTCACTGCTCCCGCTCTCCAGGAGCTTATCTTCGCCCAT   | 1207                   |
| Db         | 661   | GGGTGAGTGTACCCAGCCCTTCACTGCTCCCGCTCTCCAGGAGCTTATCTTCGCCCAT   | 720                    |
| Qy         | 1208  | CTCCCAATAAGTGGGCCCTTGTGCTGTGAGGAGACCAAGCCTCAAGGAAGATAAGAG    | 1267                   |
| Db         | 721   | CTCCCAATAAGTGGGCCCTTGTGCTGTGAGGAGACCAAGCCTCAAGGAAGATAAGAG    | 780                    |
| Qy         | 1268  | ATATGGAGATGGAGGGGGAGGACCAAGGG - CAGAGAGTAGGCTCTAGCTGGCTATCTC | 1325                   |
| Db         | 781   | TATGAAAGTGGAGGGGGAGGACCAAGGGGGCAAAAAGAGGCTCTAACTGGCTTATC     | 840                    |
| Qy         | 1326  | TGGCCTTTACTAAACACCCCCC 1345                                  |                        |
| Db         | 841   | TCITGGCCTTAACCTACACC 860                                     |                        |
| RESULT 7   |   |  |                        |
| BI819064   |   |  |                        |
| LOCUS      | 60303417F1 NIH_MGC_115 Homo sapiens   | 761 bp   | linear EST 04-OCN-2001 |
| DEFINITION | mRNA sequence.  |  |                        |
| ACCESSION  | BI819064  |  |                        |
| VERSION    | BI819064.1  | GI:15930614  |                        |
| KEYWORDS   | EST.  |  |                        |
| SOURCE     | human.  |  |                        |
| ORGANISM   | Homo sapiens  |  |                        |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |                        |
|            | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  |  |                        |
| REFERENCE  | 1 (bases 1 to 761)  |  |                        |
| AUTHORS    | NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .   |  |                        |
| TITLE      | National Institutes of Health, Mammalian Gene Collection (MGC)  |  |                        |
| JOURNAL    | Unpublished (1999)  |  |                        |
| COMMENT    | Contact: Robert Strausberg, Ph.D.<br>Email: <a href="mailto:cgapbs-re@mail.nih.gov">cgapbs-re@mail.nih.gov</a><br>Tissue Procurement: Life Technologies, Inc.<br>cDNA Library Preparation: Life Technologies, Inc.<br>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a><br>Plate: L1AM11434 row: b column: 05<br>High quality sequence stop: 759. |  |                        |

```
Db 720 CTCCTCCCTGTAATCCCTGGCTAGTACCTCTCTGCTGT 761
|||||
BM981939 702 bp mRNA linear EST 21-MAR-2002
UI-CF-EN1-adg-h-11-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adg-h-11-0-UI 3', mRNA sequence.
BM981939
BM981939.1 GI:19604935
EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 702)
JOURNAL Ronaldo, M.F., Lennon, G. and Soares, M.B.
MEDLINE Normalization and subtraction: two approaches to facilitate gene
COMMENT Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLA+Yes
FEATURES
source Location/Qualifiers
1..702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UI-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGGT.
TAG_LIB=UI-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTGCTCAGGT"
BASE COUNT 166 a 181 c 171 g 184 t
ORIGIN
Query Match 26.3%; Score 702; DB 14; Length 702;
Best Local Similarity 100.0%; Pred. No. 9.5e-69;
Matches 702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1953 ATGTGTGTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGTGGGAGACAGATTAA 2012
|||||
702 ATGTGTGTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGTGGGAGACAGATTAA 643
|||||
2013 GCAAACTCCCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCTCTGACACCC 2072
|||||
642 GCAAACTCCCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCTCTGACACCC 583
|||||
2073 ATGCTCTTTCTTTAACTTCTCAACAGATACACAGGCGCTAAACTTGTCTTACCTCCCTC 2132
|||||
582 ATGCTCTTTCTTTAACTTCTCAACAGATACACAGGCGCTAAACTTGTCTTACCTCCCTC 523
|||||
2133 CTACTGAGTCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATAT 2192
|||||
522 CTACTGAGTCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATAT 463
|||||
2193 GAGTAAAAAAGTCATGTGGGTATGTCGGGTAGAGAGAGGGGTAGCAAGTTTCATGTG 2252
|||||
462 GAGTAAAAAAGTCATGTGGGTATGTCGGGTAGAGAGAGGGGTAGCAAGTTTCATGTG 403
|||||
2253 TCCCTCTTGTGCATATATCTCCAAAGCTCTGATCCCTGCCATGGGAAAGTGGACAGAAA 2312
|||||
402 TCCCTCTTGTGCATATATCTCCAAAGCTCTGATCCCTGCCATGGGAAAGTGGACAGAAA 343
|||||
2313 CATGAGGTCATGACCTGCAGGCATCTTTACTGAGCTCTGCCGCTGGAGGGGAGAGG 2372
|||||
342 CATGAGGTCATGACCTGCAGGCATCTTTACTGAGCTCTGCCGCTGGAGGGGAGAGG 283
|||||
2373 GGGAGGAAGATGATCGCTGCACATTTCTGAGGCTACTGCTTGTCTTCAAGGCAGAA 2432
|||||
282 GGGAGGAAGATGATCGCTGCACATTTCTGAGGCTACTGCTTGTCTTCAAGGCAGAA 223
|||||
2433 ATCTTGTCTGACAGTCAGCGCTCCAGTTTGGGCGCCGATAGGAAGTTCTCCGTGGCC 2492
|||||
222 ATCTTGTCTGACAGTCAGCGCTCCAGTTTGGGCGCCGATAGGAAGTTCTCCGTGGCC 163
|||||
2493 TCCCTCAGGAGGAGGAGGCTGACATTCGAGCTCTCTCTGGGCGCCGAGGAGGAGG 2552
|||||
162 TCCCTCAGGAGGAGGAGGAGGCTGACATTCGAGCTCTCTCTGGGCGCCGAGGAGGAG 103
|||||
2553 GTTGAGGAGATCCATCCCATAGACAGCTCTGGGCTCTTGTGATTTGAGTTTTCAGAA 2612
|||||
102 GTTGAGGAGATCCATCCCATAGACAGCTCTGGGCTCTTGTGATTTGAGTTTTCAGAA 43
|||||
2613 TTAACTGCAAGTATTTGGAAAGCAAAAAAAGAAAAA 2654
|||||
42 TTAACTGCAAGTATTTGGAAAGCAAAAAAAGAAAAA 1
|||||
RESULT 9
LOCUS BG283602 1051 bp mRNA linear EST 21-FEB-2001
DEFINITION 602407930F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520191 5',
mRNA sequence.
ACCESSION BG283602
VERSION BG283602.1 GI:13033710
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1051)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csaps-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```

Plate: LLaM10417 row: k column: 08

High quality sequence stop: 7822

## FEATURES

```

FEATURES
source
Location/Qualifiers
1. .1051
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4520191"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

| BASE COUNT | 275 a  | 287 c   | 314 g | 175 t |  |
|------------|--|---|-------|-------|--|
| ORIGIN     | Query Match 26.2%; Score 701.2; DB 12; Length 1051;<br>Best Local Similarity 99.3%; Pred. No. 8 3e-69;<br>Matches 725; Conservative 0; Mismatches 3; Indels 2; Gaps 2; |   |       |       |  |
| QY         | 40   | ACATGACGGGACGCCAGGCGCGTGTGCCACCGGGATGCGAGGCCCCCGAGCGCTCCC     | 99    |       |  |
| Db         | 1  | ACATGACGGGACGCCAGGCGCGTGTGCCACCGGGATGCGAGGCCCCCGAGCGCTCCC     | 60    |       |  |
| QY         | 100  | CGCCCTCCAGTCCGAGCTAGACCTCAGCGGCRAGGTGATGCTTCTGGGAGACACAGCG    | 159   |       |  |
| Db         | 61   | CGCCCTCCAGTCCGAGCTAGACCTCAGCGGCRAGGTGATGCTTCTGGGAGACACAGCG    | 120   |       |  |
| QY         | 160  | TCGGCAAAACATGTTTCTGATCCAAATCAAAGACGGGGCTTCTCTGTCGGAACCTTCA    | 219   |       |  |
| Db         | 121  | TCGGCAAAACATGTTTCTGATCCAAATCAAAGACGGGGCTTCTCTGTCGGAACCTTCA    | 180   |       |  |
| QY         | 220  | TAGCCACCGTGGGATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGA    | 279   |       |  |
| Db         | 181  | TAGCCACCGTGGGATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGA    | 240   |       |  |
| QY         | 280  | AGCTGTCAGATCTGGGACACCGCTGGGCAGGACGGTTCCGGAAGCGTCAACCCATGCTTAT | 339   |       |  |
| Db         | 241  | AGCTGTCAGATCTGGGACACCGCTGGGCAGGACGGTTCCGGAAGCGTCAACCCATGCTTAT | 300   |       |  |
| QY         | 340  | ACAGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATCACCAAAATCTCTTTTCGACA   | 399   |       |  |
| Db         | 301  | ACAGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATCACCAAAATCTCTTTTCGACA   | 360   |       |  |
| QY         | 400  | ACATCAGGGCGTGGCTCACTAGATTCATGAGTATGCCAGAGGGACGTGGTGATCATGC    | 459   |       |  |
| Db         | 361  | ACATCAGGGCGTGGCTCACTAGATTCATGAGTATGCCAGAGGGACGTGGTGATCATGC    | 420   |       |  |
| QY         | 460  | TGCTAGCAACAGGCGGATATGACGACGAAAGAGTGATCCGTCCGAGACGCGAGAGA      | 519   |       |  |
| Db         | 421  | TGCTAGCAACAGGCGGATATGACGACGAAAGAGTGATCCGTCCGAGACGCGAGAGA      | 480   |       |  |
| QY         | 520  | CCATTGGCCAGGAGTACGGTCTCCCTTCCTTGGAGACGACGCCCAAGACTGGCATGAATG  | 579   |       |  |
| Db         | 481  | CCATTGGCCAGGAGTACGGTCTCCCTTCCTTGGAGACGACGCCCAAGACTGGCATGAATG  | 540   |       |  |
| QY         | 580  | TGGAGTTAGCCTTTCTGGCCATGCCAAGGAACGTGAAATACCGGGCCGGGCATCAGCGGG  | 639   |       |  |
| Db         | 541  | TGGAGTTAGCCTTTCTGGCCATGCCAAGGAACGTGAAATACCGGGCCGGGCATCAGCGGG  | 600   |       |  |
| QY         | 640  | ATGAGCCACGCTCCAGATCCGAGACTATGTAGAGTCCCAG-AGAGAGCGCTCCAGCTGC   | 698   |       |  |
| Db         | 601  | ATGAGCCACGCTCCAGATCCGAGACTATGTAGAGTCCCAG-AGAGAGCGCTCCAGCTGC   | 660   |       |  |
| QY         | 699  | TGCTCTCTCATGTGAATCCAGGGG-CAGAGAGGAGGCTCTGGAGGCACACAGAGTGA     | 757   |       |  |
| Db         | 661  | TGCTCTCTCATGTGAATCCAGGGG-CAGAGAGGAGGCTCTGGAGGCACACAGAGTGA     | 720   |       |  |
| QY         | 758  | GCCTTCCCC 767   |       |       |  |
| Db         | 721  | GCCTTCCCC 730   |       |       |  |





```

Db      61  AGGCATGCAGCTCCACAGCGGATTTCTTCAAGCAGCTGAAGTGTAGCCCTCCTGGGTT 120
QY      1914  AAGAGCCAGATAGAGGAGAAATCCCTTCTAGGTTTGAATGTGTTGTGAAGAAAAAGAG 1973
Db      121  AAGAGCCAGATAGAGGAGAAATCCCTTCTAGGTTTGAATGTGTTGTGAAGAAAAAGAG 180
QY      1974  AAATCCCTGGCTCCTGGAGCTGGTGGGACACAGATTAAGCAAACTCCCTGCACATGTA 2033
Db      181  AAATCCCTGGCTCCTGGAGCTGGTGGGACACAGATTAAGCAAACTCCCTGCACATGTA 240
QY      2034  TCCCTTTGACCCCAAGCTCTGCTCTCCTCTGACCAACCCATGCCCTTCTTTAACTTCT 2093
Db      241  TCCCTTTGACCCCAAGCTCTGCTCTCCTCTGACCAACCCATGCCCTTCTTTAACTTCT 300
QY      2094  CAAACAGATACAGAGGCCCTAAACTGTTTACCTCCCTCTCTACTAGTACAGTTAGGTGG 2153
Db      301  CAAACAGATACAGAGGCCCTAAACTGTTTACCTCCCTCTCTACTAGTACAGTTAGGTGG 360
QY      2154  TGGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAAGTCATGTGG 2213
Db      361  TGGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAAGTCATGTGG 420
QY      2214  GTATGCTCTGGGTAGAGAGGGGTAGCAAGTTTCATGTGTCTCTTGTTCACATATCTC 2273
Db      421  GTATGCTCTGGGTAGAGAGGGGTAGCAAGTTTCATGTGTCTCTTGTTCACATATCTC 480
QY      2274  CCAAGCTCTGATCCCTGCCATGGGAAAGTGGACAGGAACATAGAGTCAATGACCTCAGG 2333
Db      481  CCAAGCTCTGATCCCTGCCATGGGAAAGTGGACAGGAACATAGAGTCAATGACCTCAGG 540
QY      2334  CATCTTTACTGCAGCTCTCCGCCCTGGAGGGGAGAGGGGAGGAAGATATGGCGTG 2393
Db      541  CATCTTTACTGCAGCTCTCCGCCCTGGAGGGGAGAGGGGAGGAAGATATGGCGTG 600
QY      2394  CACATTT 2400
Db      601  CACATTT 607

```

```

RESULT 14
BI767046
LOCUS      603054267F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:520360 5',
DEFINITION mRNA sequence.
ACCESSION BI767046
VERSION    BI767046.1 GI:15758624
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      NIH-MGC http://mgi.nci.nih.gov/
JOURNAL    National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT    Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: c9abps-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://Image.lnl.gov
            Plate: LLAM11510 row: f column: 17
            High quality sequence stop: 707.
            Location/Qualifiers
                1..740
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="520360"
                /clone_lib="NIH_MGC_122"
                /lab_host="DH10B"

```

## FEATURES

Source

RESULT 15  
BG253976  
LOCUS

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036. Note: this is a NIH\_MGC Library."

```

BASE COUNT      171 a      198 c      222 g      149 t
ORIGIN
Query Match      22.5%; Score 601.6; DB 13; Length 740;
Best Local Similarity 96.2%; Pred. No. 9.8e-58;
Matches 712; Conservative 0; Mismatches 19; Indels 9; Gaps 9;
QY      42  ATGACCGGACCGGCGGCGGTTGCCACCGGATGGCGAGGCCCGGCGCTCCCG 101
Db      1  ATGACCGGACCGGCGGCGGTTGCCACCGGATGGCGAGGCCCGGCGCTCCCG 60
QY      102  CCCGTCAGTCC -GAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGT 160
Db      61  CCCGTCAGTCTCTGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGT 120
QY      161  CGGCAAAACATGTTTCTGATCCAAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCAT 220
Db      121  CGGCAAAACATGTTTCTGATCCAAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCAT 180
QY      221  AGCACCCTCGGATAGACTTCAGGAACAGGTGGTGACTGTGGATGGCGTGAAGTAA 280
Db      181  AGCACCCTCGGATAGACTTCAGGAACAGGTGGTGACTGTGGATGGCGTGAAGTAA 240
QY      281  GCTGTCAGATCTGGGACACCTCGGCAAGGAACTCCGAAGCGTCAACCATGCTTATTA 340
Db      241  GCTGTCAGATCTGGGACACCTCGGCAAGGAACTCCGAAGCGTCAACCATGCTTATTA 300
QY      341  CAGATGTCAGGCTTGTCTGTGTATGATACATCACCACAAATCTTCTTTTCGACAA 400
Db      301  CAGATGTCAGGCTTGTCTGTGTATGATACATCACCACAAATCTTCTTTTCGACAA 360
QY      401  CATCAGGCTCGCTCAGTTCATGATTCATGATTCGCCAGGAGGACGTGG -TGATCATGC 459
Db      361  CATCAGGCTCGCTCAGTTCATGATTCATGATTCGCCAGGAGGACGTGGTGTGATATGC 420
QY      460  TGCTAGGCAACAAAGCGGATATGAGCAGCAAGAGTGTATCCGTTCCGAAGACGGAGAGA 519
Db      421  TGCTAGGCAACAAAGCGGATATGAGCAGCAAGAGTGTATCCGTTCCGAAGACGGAGAGA 480
QY      520  CTTTGGCCAGGAGTACGGTGTCTTCCCTTCT -GGAGACCAAGCGCCAAAGAC -TGTCATGAA 577
Db      481  CTTTGGCCAGGAGTACGGTGTCTTCCCTTCTTCCCTGGGAGAACAGTGTCCAAAGACTTGGCATGAA 540
QY      578  TGTGGAGTAA -GCCTTCTGCGCATCGCCAGGAACCTGAATACCGGCGCGGCGCATCAGG 636
Db      541  TGTGGAGTAAAGGCTTTTCTTGGCATCGCCAGGAACCTGAATACCGGCGCGGCGCATCAGG 600
QY      637  CGGATGAGCCAGCTTCCAGA -TCCGAGACTATGTAG -AGTCCCAAGAAAGCGCTCCAG 694
Db      601  CGGATGAACCCAGCTTCCAGATTCGAGACTATGTAGAAGTCCCAAGAAAGCGCTCAG 660
QY      695  C-TGCTGCTCCTTCATGTGAA -TCCAGGGGGCAGAGAGGAGGCTCTGGAGGACACACAGG 752
Db      661  CTTGCTGCTCCTTCATGTGAAGTCCCGAGGGGCGAGAGAGGAGGCTCTGGAGCGCCACAGG 720
QY      753  ATGAGCCTTCCCTCCCA 772
Db      721  ATGAGCCTTCTTCCATCCA 740

```

BG253976 691 bp mRNA linear EST 13-FEB-2001

```
DEFINITION 602366910F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475067 5',
ACCESSION BG253976
VERSION BG253976.1 GI:12763792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10300 row: c column: 04
High quality sequence stop: 650.
FEATURES
    Location/Qualifiers
        1..691
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4475067"
            /clone_lib="NIH_MGC_91"
            /tissue_type="adenocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: prostate; Vector: pCMV-SPORT6; Site.1: NotI;
            Site.2: SalI; Cloned unidirectionally; oligo-dr primed.
            Average insert size 1.4 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
BASE COUNT 175 a 185 c 190 g 141 t
ORIGIN
Query Match 22.5%; Score 600.6; DB 12; Length 691;
Best Local Similarity 97.6%; Pred. No. 1.3e-57;
Matches 663; Conservative 0; Mismatches 9; Indels 7; Gaps 5;
QY 287 GATCTGGGACACCGCTGGGAGGACGGTTCGGAGGCTACCCATGCTTATTACAGAGA 346
DB 14 GATCTGGGACACCGCTGGGAGGACGGTTCGGAGGCTACCCATGCTTATTACAGAGA 73
QY 347 TGCTCAGGCTTGCTTCGCTGTATGACATCACCAACAAATCTTCTTCGACACATCAG 406
DB 74 TGCTCAGGCTTGCTTCGCTGTATGACATCACCAACAAATCTTCTTCGACACATCAG 133
QY 407 GGCCTGGCTCACTGAGATTCATGATGCCAGAGGACGCTGGTATCATGCTGCTAGG 466
DB 134 GGCCTGGCTCACTGAGATTCATGATGCCAGAGGACGCTGGTATCATGCTGCTAGG 193
QY 467 CAACAAGCGGATATGACGACGCGAAGAGTGATCCGTTCCGAAGACGGAGACCTTGGC 526
DB 194 CAACAAGCGGATATGACGACGCGAAGAGTGATCCGTTCCGAAGACGGAGACCTTGGC 253
QY 527 CAGGGAGTACGGTTCCTTCCTGGAGACGACGCGCAAGACTGGCATGAATGTGGAGTT 586
DB 254 CAGGGAGTACGGTTCCTTCCTGGAGACGACGCGCAAGACTGGCATGAATGTGGAGTT 313
QY 587 AGCCTTTCTGGCCATCGCAAGGAACCTGAATACCGGCGCGGATCAGGCGGATGAGCC 646
DB 314 AGCCTTTCTGGCCATCGCAAGGAACCTGAATACCGGCGCGGATCAGGCGGATGAGCC 373
QY 647 CAGCTTCAGATCCGAGACTATGTAGTCCGAGAGAGCGCTCCAGCTGCTGCTCCT 706
DB 374 CAGCTTCAGATCCGAGACTATGTAGTCCGAGAGAGCGCTCCAGCTGCTGCTCCT 433
QY 707 CATGTGAATCCAGGGGACAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 766
|||||
434 CATGTGAATCCAGGGGACAGAGGAGGCTCTGGAGGCACACAGGATGCAG-CTTCCCC 492
767 CTCCAGGCTGGCTTATTCAGAGGCTGAGCAATGGGA--GAAAGATGGAGGACTC 824
493 CTCCAGGCTGGCTTATTCAGAGGCTGAGCAATGGGAACGATGGAGGACTC 552
825 ACTGCACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTTACTTGTGCTGCG-GG 883
553 ACTGCACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTTACTTGTGCTGCGAGG 612
884 TCTCCCGCATCCACAGGGA--GGTAAACACTAGCTTTTATTTT--AATAGTACATAA 940
613 TCTCCCGCATCCACAGGAGGAGGTTATTTCTCAATACGTACATAA 672
941 TTTAATACCAAAAAAGGCG 959
673 TTTAATACCAAAAAAGGCG 691
Search completed: June 24, 2003, 00:50:48
Job time : 3651 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:35:29 ; Search time 121 Seconds  
(without alignments)  
6777.303 Million cell updates/sec

Title: US-09-817-199B-1  
Perfect score: 2674  
Sequence: 1 ttccgctggggccggcact.....aaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2607.8 | 97.5        | 2612   | 4     | US-09-484-970B-142 |
| 2          | 873    | 32.6        | 875    | 4     | US-09-075-454-10   |
| 3          | 292.6  | 10.9        | 1340   | 2     | US-08-824-873-2    |
| 4          | 292.6  | 10.9        | 1340   | 3     | US-09-198-184-2    |
| 5          | 193.4  | 7.2         | 99500  | 4     | US-09-798-096-10   |
| 6          | 193.2  | 7.2         | 87350  | 3     | US-08-781-891-79   |
| 7          | 193.2  | 7.2         | 87543  | 4     | US-09-791-211-3    |
| 8          | 192.8  | 7.2         | 8453   | 4     | US-09-167-681-45   |
| 9          | 192.4  | 7.2         | 112132 | 4     | US-09-741-150-3    |
| 10         | 192.2  | 7.2         | 6678   | 3     | US-08-816-617B-1   |
| 11         | 192.2  | 7.2         | 9301   | 4     | US-09-449-218D-18  |
| 12         | 192.2  | 7.2         | 35060  | 3     | US-08-814-095-7    |
| 13         | 192.2  | 7.2         | 98844  | 4     | US-09-791-211-10   |
| 14         | 191.4  | 7.2         | 29629  | 4     | US-09-729-995-3    |
| 15         | 190.8  | 7.1         | 6769   | 1     | US-08-480-784-20   |
| 16         | 190.8  | 7.1         | 6769   | 1     | US-08-483-553-20   |
| 17         | 190.8  | 7.1         | 6769   | 1     | US-08-487-002-20   |
| 18         | 190.8  | 7.1         | 6769   | 1     | US-08-483-554B-20  |
| 19         | 190.8  | 7.1         | 6769   | 1     | US-08-488-011B-20  |
| 20         | 190.8  | 7.1         | 6769   | 4     | US-08-850-727-20   |
| 21         | 190.8  | 7.1         | 6769   | 5     | PCT-US95-10202-20  |
| 22         | 190.8  | 7.1         | 6769   | 5     | PCT-US95-10203-20  |
| 23         | 190.8  | 7.1         | 6769   | 5     | PCT-US95-10220-20  |
| 24         | 190.4  | 7.1         | 45546  | 4     | US-09-146-053-6    |
| 25         | 190    | 7.1         | 45716  | 4     | US-08-965-048-5    |
| 26         | 190    | 7.1         | 45989  | 4     | US-08-965-048-6    |
| 27         | 189    | 7.1         | 14581  | 4     | US-08-520-373D-4   |

|    |       |     |       |   |                   |                   |
|----|-------|-----|-------|---|-------------------|-------------------|
| 28 | 189   | 7.1 | 22481 | 4 | US-08-367-841A-43 | Sequence 43, Appl |
| 29 | 189   | 7.1 | 22481 | 5 | PCT-US95-07201-43 | Sequence 43, Appl |
| 30 | 189   | 7.1 | 22484 | 4 | US-09-875-223-2   | Sequence 2, Appl1 |
| 31 | 189   | 7.1 | 55827 | 4 | US-09-813-133A-3  | Sequence 3, Appl1 |
| 32 | 189   | 7.1 | 99500 | 4 | US-09-798-096-10  | Sequence 10, Appl |
| 33 | 188.4 | 7.0 | 36651 | 4 | US-09-738-894A-3  | Sequence 3, Appl1 |
| 34 | 188.4 | 7.0 | 59085 | 4 | US-09-813-817-3   | Sequence 3, Appl1 |
| 35 | 188.4 | 7.0 | 59085 | 4 | US-09-978-197-3   | Sequence 3, Appl1 |
| 36 | 187.8 | 7.0 | 50000 | 4 | US-09-146-053-3   | Sequence 3, Appl1 |
| 37 | 187.6 | 7.0 | 8133  | 4 | US-09-659-791A-10 | Sequence 10, Appl |
| 38 | 187   | 7.0 | 3867  | 4 | US-09-347-114A-81 | Sequence 81, Appl |
| 39 | 187   | 7.0 | 43950 | 4 | US-09-735-934A-3  | Sequence 3, Appl1 |
| 40 | 186.6 | 7.0 | 8353  | 3 | US-08-611-587-1   | Sequence 1, Appl1 |
| 41 | 186.4 | 7.0 | 1043  | 4 | US-09-165-868-4   | Sequence 4, Appl1 |
| 42 | 186.4 | 7.0 | 3694  | 4 | US-09-232-200-46  | Sequence 46, Appl |
| 43 | 186.4 | 7.0 | 3694  | 4 | US-09-232-197-46  | Sequence 46, Appl |
| 44 | 186.4 | 7.0 | 3694  | 4 | US-09-232-201-46  | Sequence 46, Appl |
| 45 | 186.4 | 7.0 | 3704  | 4 | US-09-232-200-24  | Sequence 24, Appl |

## ALIGNMENTS

## RESULT 1

US-09-484-970B-142  
; Sequence 142, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484, 970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 142  
; LENGTH: 2612  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1  
US-09-484-970B-142

|                       |              |   |            |              |
|-----------------------|--------------|---|------------|--------------|
| Query Match           | 97.5%        | Score 2607.8;   | DB 4;      | Length 2612; |
| Best Local Similarity | 99.9%        | Pred. No. 0;  |            |              |
| Matches 2609;         | Conservative | 0;  | Mismatches | 2;           |
|                       |              |   | Indels     | 0;           |
|                       |              |   | Gaps       | 0;           |
| <hr/>                 |              |   |            |              |
| QY                    | 27           | CTCTCTCCAGGACATGACGGGACGCCGCGGTCAGGCGCCGTTGCCACCCGGGATGGCGAGGCC | 86         |              |
| Db                    | 1            | CTCTCTCCAGGACATGACGGGACGCCGCGGTCAGGCGCCGTTGCCACCCGGGATGGCGAGGCC | 60         |              |
| <hr/>                 |              |   |            |              |
| QY                    | 87           | CCGAGGGCTCCCGCCCTGCAGTCCGAGCTAGACCTCAGGGCAAGGTGATGCTCTG         | 145        |              |
| Db                    | 61           | CCGAGGGCTCCCGCCCTGCAGTCCGAGCTAGACCTCAGGGCAAGGTGATGCTCTG         | 120        |              |
| <hr/>                 |              |   |            |              |
| QY                    | 147          | GGAGACACAGCGCTCGGCAAAACATGTTCTCTGATCAATTCAGGCGGCTTCCTG          | 206        |              |
| Db                    | 121          | GGAGACACAGCGCTCGGCAAAACATGTTCTCTGATCAATTCAGGCGGCTTCCTG          | 180        |              |
| <hr/>                 |              |   |            |              |
| QY                    | 207          | TCCGGAACCTTCATAGCCACCGTCGGGATAGATTCAGGAACAAGGTGCTGATGAT         | 266        |              |
| Db                    | 181          | TCCGGAACCTTCATAGCCACCGTCGGGATAGATTCAGGAACAAGGTGCTGATGAT         | 240        |              |
| <hr/>                 |              |   |            |              |
| QY                    | 267          | GGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT         | 326        |              |
| Db                    | 241          | GGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT         | 300        |              |
| <hr/>                 |              |   |            |              |
| QY                    | 327          | ACCATGCTTATTACAGAGATGCTCAGGCGCTTCCTCTGCTGTATGACATCACCACAA       | 386        |              |
| Db                    | 301          | ACCATGCTTATTACAGAGATGCTCAGGCGCTTCCTCTGCTGTATGACATCACCACAA       | 360        |              |

QY 387 TCCTCTTCGACACATCAGGGCTGGCTCACTGAGATTATGAGTATGCCACAGGGAC 446  
DB 361 TCCTCTTCGACACATCAGGGCTGGCTCACTGAGATTATGAGTATGCCACAGGGAC 420  
QY 447 GTGGTGATCATGCTGCTAGGCAACAAGCGGATATGAGCAGCAAGAGATGATCCGTTCC 506  
DB 421 GTGGTGATCATGCTGCTAGGCAACAAGCGGATATGAGCAGCAAGAGATGATCCGTTCC 480  
QY 507 GAAGACGGAGAGACTTGGCCAGGAGTACGGTGTCCCTTCTCGAGACAGCGCAAG 566  
DB 481 GAAGACGGAGAGACTTGGCCAGGAGTACGGTGTCCCTTCTCGAGACAGCGCAAG 540  
QY 567 ACTGGCATGATGTGGAGTTAGCCCTTCTGGCCATCGCAAGGAACCTGAATATCCGGGCC 626  
DB 541 ACTGGCATGATGTGGAGTTAGCCCTTCTGGCCATCGCAAGGAACCTGAATATCCGGGCC 600  
QY 627 GGGCATCAGGCGGATGAGCCAGCTTCCAGATCCAGACTATGTAGAGTCCAGAAAG 686  
DB 601 GGGCATCAGGCGGATGAGCCAGCTTCCAGATCCAGACTATGTAGAGTCCAGAAAG 660  
QY 687 CGCTCAGCTGCTCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCA 746  
DB 661 CGCTCAGCTGCTCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCA 720  
QY 747 CACAGGATGAGCCCTCCCTCCAGGCTTCTAGCAGGGAGCTTATCTCCAACCTCC 806  
DB 721 CACAGGATGAGCCCTCCCTCCAGGCTTCTAGCAGGGAGCTTATCTCCAACCTCC 780  
QY 807 GAGAAAGATGAGGACTCACTGACAGCCGCTTCTAGCAGGGAGCTTATCTCCAACCTCC 866  
DB 781 GAGAAAGATGAGGACTCACTGACAGCCGCTTCTAGCAGGGAGCTTATCTCCAACCTCC 840  
QY 867 TACTTGTAGTCTCCGGTCTCCCGATCCACAGGGAGGTAAACACTTAGCTTTTATT 926  
DB 841 TACTTGTAGTCTCCGGTCTCCCGATCCACAGGGAGGTAAACACTTAGCTTTTATT 900  
QY 927 TTAATAGTACATAATTTAATACCAAAAAGCGCTGATCCCAAAAACCGAGGCTGG 986  
DB 901 TTAATAGTACATAATTTAATACCAAAAAGCGCTGATCCCAAAAACCGAGGCTGG 960  
QY 987 GAGTACTGCGCCCTTTGCTTCTAGGACTTGGGGGCGGCGCTCCCTCTTAAGCATAA 1046  
DB 961 GAGTACTGCGCCCTTTGCTTCTAGGACTTGGGGGCGGCGCTCCCTCTTAAGCATAA 1020  
QY 1047 CAAGGTGGTGTGCTCCAGCTCAGCCCGCAGGGGACACAGATGACTTTGGGGTGAGG 1106  
DB 1021 CAAGGTGGTGTGCTCCAGCTCAGCCCGCAGGGGACACAGATGACTTTGGGGTGAGG 1080  
QY 1107 CAGGTAATGACTCCATCGCACCCCTCAGTTTCCAGCTGGACAGGCTCAGGTGACCCAGCC 1166  
DB 1081 CAGGTAATGACTCCATCGCACCCCTCAGTTTCCAGCTGGACAGGCTCAGGTGACCCAGCC 1140  
QY 1167 TTCCTCTCTCCGCTCTCCAGGAGCTTATCTTGGCCCATCTCCCAAAATAGTGGGCC 1226  
DB 1141 TTCCTCTCTCCGCTCTCCAGGAGCTTATCTTGGCCCATCTCCCAAAATAGTGGGCC 1200  
QY 1227 TTGTGCTGTAGGAGACCAAGCCCTCAGGAGAGATAAGAGATGAGAGATGGAGGGG 1286  
DB 1201 TTGTGCTGTAGGAGACCAAGCCCTCAGGAGAGATAAGAGATGAGAGATGGAGGGG 1260  
QY 1287 AGGACAGGGGACAGAGATGAGGTCTAGCTGGCTTATCTCTGGCTTACTAACACCCCT 1346  
DB 1261 AGGACAGGGGACAGAGATGAGGTCTAGCTGGCTTATCTCTGGCTTACTAACACCCCT 1320  
QY 1347 GGAGGATGCCCTTTCTCCAGCACACAAGCATTGGGGCCTCTGGAATATTTGGTTC 1406  
DB 1321 GGAGGATGCCCTTTCTCCAGCACACAAGCATTGGGGCCTCTGGAATATTTGGTTC 1380  
QY 1407 CAGGCTCCTGTTCTGAGCTTCAGATCTCGGGGAGCCCTCCCGCCCTGATCCCTG 1466  
DB 1381 CAGGCTCCTGTTCTGAGCTTCAGATCTCGGGGAGCCCTCCCGCCCTGATCCCTG 1440

QY 1467 GCTTAGCTACCTTCCTGCTGTGCACCTAAACACTCAGGTCAAGACTAGAAAAGATT 1526  
DB 1441 GCTTAGCTACCTTCCTGCTGTGCACCTAAACACTCAGGTCAAGACTAGAAAAGATT 1500  
QY 1527 TTGTTTTTATTTTTTGAATGGAGTCTCGTCTGTGCGCCAGGCTGAGGTGCGAGTAGTG 1586  
DB 1501 TTGTTTTTATTTTTTGAATGGAGTCTCGTCTGTGCGCCAGGCTGAGGTGCGAGTAGTG 1560  
QY 1587 CAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCTCCACCTCAGCC 1646  
DB 1561 CAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCTCCACCTCAGCC 1620  
QY 1647 GCCAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGCTAAATTTTGTATTTTTT 1706  
DB 1621 CCCAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGCTAAATTTTGTATTTTTT 1680  
QY 1707 GTAGACACAGGTTTCCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCNA 1766  
DB 1681 GTAGACACAGGTTTCCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCNA 1740  
QY 1767 CTGCGGCTCGGCTCCCAAGTACTGGGATTACAGCAAGAGCCACCATGCCAGGCT 1826  
DB 1741 CTGCGGCTCGGCTCCCAAGTACTGGGATTACAGCAAGAGCCACCATGCCAGGCT 1800  
QY 1827 AGATGTCTTATCCCAATCCTTTGGCAGGATCGAGTCCACAGGCGATTTCTTCAAGC 1886  
DB 1801 AGATGTCTTATCCCAATCCTTTGGCAGGATCGAGTCCACAGGCGATTTCTTCAAGC 1860  
QY 1887 AGCTGAAGTGTGTTAGCCCTCCTGGTTAAGAGCCAGATAGAGGAATCCCTTCCTAGG 1946  
DB 1861 AGCTGAAGTGTGTTAGCCCTCCTGGTTAAGAGCCAGATAGAGGAATCCCTTCCTAGG 1920  
QY 1947 TTTGGAATGTGTTGAAAAAAGAGAAATCCCTGGCTCCTGAGCTGTTGGGAGACAA 2006  
DB 1921 TTTGGAATGTGTTGAAAAAAGAGAAATCCCTGGCTCCTGAGCTGTTGGGAGACAA 1980  
QY 2007 GATTAACCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTGA 2066  
DB 1981 GATTAACCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTGA 2040  
QY 2067 CCACCATGCCCTTTCCCTTTAACTTCTCAACAGATACAGGGCTTAACTGCTTACCT 2126  
DB 2041 CCACCATGCCCTTTCCCTTTAACTTCTCAACAGATACAGGGCTTAACTGCTTACCT 2100  
QY 2127 CCCCTCTACTGAGTCAAGTGTGTTGAGGAGTCAACCATTTCCGAGTTAAACCAATG 2186  
DB 2101 CCCCTCTACTGAGTCAAGTGTGTTGAGGAGTCAACCATTTCCGAGTTAAACCAATG 2160  
QY 2187 CAATATCAGTAAACAAAGTCATGTGGTATGTCTGGGTAGAGAGGGGTAGCAAGTT 2246  
DB 2161 CAATATCAGTAAACAAAGTCATGTGGTATGTCTGGGTAGAGAGGGGTAGCAAGTT 2220  
QY 2247 CATGTGCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAGAGTGAC 2306  
DB 2221 CATGTGCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAGAGTGAC 2280  
QY 2307 AGGAACATGAGTCACTGACCTCGAGCATCTTTTACTGCACTCTGCCGCTGGAGGG 2366  
DB 2281 AGGAACATGAGTCACTGACCTCGAGCATCTTTTACTGCACTCTGCCGCTGGAGGG 2340  
QY 2367 GAGAGGGGAGGAGAGTATGCCCTGCCATTTCTCAGGCTACTGATTTGCTTCAAG 2426  
DB 2341 GAGAGGGGAGGAGAGTATGCCCTGCCATTTCTCAGGCTACTGATTTGCTTCAAG 2400  
QY 2427 GCAGAAATCTTGCTCTGAGCAGTCAGGGCTCCAGTTTGGCCCGATAGGAAGTTCTCC 2486  
DB 2401 GCAGAAATCTTGCTCTGAGCAGTCAGGGCTCCAGTTTGGCCCGATAGGAAGTTCTCC 2460  
QY 2487 GTGGCTCTCTCAGGACAGAGGAGGAGGCTGACATTCGCCAGTCTCTTCTGGGGCCCA 2546  
DB 2461 GTGGCTCTCTCAGGACAGAGGAGGAGGCTGACATTCGCCAGTCTCTTCTGGGGCCCA 2520  
QY 2547 AGCAGGTTCCAGGAGATCCAAATCCCATACAGCTCTGGGCTCTTTCATTTGAGTTT 2606

|||||  
Db 2521 AGGAGGTTGAGAGATCCATCCCATAGACAGCTCGGGCTCTTGCAATTGAGTTT 2580  
QY 2607 TCAGATAAAGTCCAGTATTTTGGAAAGCA 2637  
|||||  
Db 2581 TCAGATAAAGTCCAGTATTTTGGAAAGCA 2611  
|||||

## RESULT 2

US-09-075-454-10  
; Sequence 10, Application US/09075454  
; Patent No. 6391580  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Lal, Preeti  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Batra, Sajeev  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: RAS PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,454  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/766,551  
; FILING DATE: DECEMBER 12, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0169-1 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 875 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: UCMLST01  
; CLONE: 1528559

US-09-075-454-10

Query Match 32.6%; Score 873; DB 4; Length 875;  
Best Local Similarity 100.0%; Pred. No. 1.5e-186;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 46 CGGGACGCGCGGTTGACCCCGGATGGCGAGCGCCCGGAGCGCTCCCGCCCT 105  
Db 1 CGGGACGCGCGGTTGACCCCGGATGGCGAGCGCCCGGAGCGCTCCCGCCCT 60  
QY 106 GCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACAGAGGCGTCGGCA 165  
Db 61 GCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACAGAGGCGTCGGCA 120

QY 166 AAACATGTTTCTGATCCAAATTCAAAGACGGGGCTTCTGTCGGAACTTCATAGCCA 225  
Db 121 AAACATGTTTCTGATCCAAATTCAAAGACGGGGCTTCTGTCGGAACTTCATAGCCA 180  
QY 226 CCGTGGCATAGACTTCAGGAAACAGGTGTGACTGTGGATGGCGTGGAGTGAAGCTGC 285  
Db 181 CCGTGGCATAGACTTCAGGAAACAGGTGTGACTGTGGATGGCGTGGAGTGAAGCTGC 240  
QY 286 AGATCTGGGACACCGCTGGGAGGAAAGGTTCCGAAGCGTCCACCATGCTTATACAGAG 345  
Db 241 AGATCTGGGACACCGCTGGGAGGAAAGGTTCCGAAGCGTCCACCATGCTTATACAGAG 300  
QY 346 ATGCTCAGGCTTGTCTTCTGCTGATGACATCACCAACAATCTTCTTCGACAACTCA 405  
Db 301 ATGCTCAGGCTTGTCTTCTGCTGATGACATCACCAACAATCTTCTTCGACAACTCA 360  
QY 406 GGGCTTGGCTCACTGAGATTTCATGATATGCCAGAGGACGCTGGTGTGATCATCTGCTAG 465  
Db 361 GGGCTTGGCTCACTGAGATTTCATGATATGCCAGAGGACGCTGGTGTGATCATCTGCTAG 420  
QY 466 GCAACAAGGCGGATATGACAGGAAAGAGTGTATCGTTCGGAAGACGAGAGACCTTGG 525  
Db 421 GCAACAAGGCGGATATGACAGGAAAGAGTGTATCGTTCGGAAGACGAGAGACCTTGG 480  
QY 526 CCAGGAGTACGGTGTCTTCTTCTGAGACACGCGCCAGAGCTGCATGAATGTGAGT 585  
Db 481 CCAGGAGTACGGTGTCTTCTTCTGAGACACGCGCCAGAGCTGCATGAATGTGAGT 540  
QY 586 TAGCCTTTCTGGCCATCGCCAAAGAACTGAAATACCGGGCGGCGCATCAGCGGATGAGC 645  
Db 541 TAGCCTTTCTGGCCATCGCCAAAGAACTGAAATACCGGGCGGCGCATCAGCGGATGAGC 600  
QY 646 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCAGAAAGAGCTCCAGCTTGTGCTCT 705  
Db 601 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCAGAAAGAGCTCCAGCTTGTGCTCT 660  
QY 706 TCATGTGAATCCAGGGGCGAGAGAGGCTCTGAGGACACACAGATGCACGCTTCCC 765  
Db 661 TCATGTGAATCCAGGGGCGAGAGAGGCTCTGAGGACACACAGATGCACGCTTCCC 720  
QY 766 CTCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGGAGACTCA 825  
Db 721 CTCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGGAGACTCA 780  
QY 826 CTGCACAGCGCTTCTTAGCAGGAGCTATATCTCAACTCTTACTTGAGTTCTCGGCTC 885  
Db 781 CTGCACAGCGCTTCTTAGCAGGAGCTATATCTCAACTCTTACTTGAGTTCTCGGCTC 840  
QY 886 TCCCGCATCCACAGGAGGCTAAACACTTAG 918  
Db 841 TCCCGCATCCACAGGAGGCTAAACACTTAG 873

## RESULT 3

US-08-824-873-2  
; Sequence 2, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS





```

; LENGTH: 87543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 7421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7427
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 11609
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12605
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12742
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29370
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29979
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29980
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29981
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30136
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30140
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31205
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31206
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31592
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33095
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33160
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34066
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34072
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 36816
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 39020
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42164
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42459
; OTHER INFORMATION: unknown

```

```

; NAME/KEY: unsure
; LOCATION: 46808
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46823
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46826
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 47291
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52786
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52787
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 53384
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 54684
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59215
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59235
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59242
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 63290
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 66614
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68660
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68697
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68718
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68733
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68739
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 69785
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 79134
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 79198
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 86336
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
; US-09-791-211-3

```

```

Query Match 7.2%; Score 193.2; DB 4; Length 87543;
Best Local Similarity 79.9%; Pred. No. 1e-33;
Matches 239; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

```

QY 1525 TTTTGTGTTTATTTTGAATGGAGTCTCGTCTGTGTCGCCAGGCTGAGTGCAGTAG 1584  
Db 42476 TTTTGTGTTTATTTTNGAGACAGAGTCTCGTCTGTGTCGCCAGGCTGAAGTGCAGTGG 42417  
QY 1585 TGCATCTCGCTCACTACAACTCCACCTCCCTGGGCTCAAGCGATCCCTCCACCTCAG 1644  
Db 42416 CACAATCTGACCTCACTGCACTCCACCTCCT-GGTTCAAGCAATCTCTCGCTCAG 42358  
QY 1645 CCGCCAAGTAGCTGGGACTATAGGTGTGTACCATCACACTGGCTAAATTTTGTATTTT 1704  
Db 42357 CTTCTGAAATAGTGGGATTTACAGGACCAACACCAAGCCAGCTAAATTTTGTACTTT 42298  
QY 1705 TTCTGACACAGGTTTCCGCTGTCGCCAGGCTGGTCTGTAATCTCTGAGTCAAGCA 1764  
Db 42297 TAGTAAAGCGGATTTTCATCATGTTGGCCAGGCTGCTGTGAATCTCTGACCTCAAGT 42238  
QY 1765 ACCTGCGGCTCGGCTCCCAAGTACTGGGATTACACGAAAGGACCATGCCAG 1823  
Db 42237 ATCTGCTGCTCGGCTCCCAAGTGTGGGATTACAGGCTGAGCCACCATGCCAG 42179

## RESULT 8

US-09-167-681-45/c  
; Sequence 45, Application US/09167681A  
; Patent No. 6265561  
; GENERAL INFORMATION:  
; APPLICANT: Weinshilboum, M.D., Richard M.  
; APPLICANT: Raftogiannis, Rebecca B.  
; APPLICANT: Wood, Thomas C.  
; APPLICANT: Ottewill, Diane M.  
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
; FILE REFERENCE: 07039/118001  
; CURRENT APPLICATION NUMBER: US/09/167,681A  
; CURRENT FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 8447  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (4361)...(4507)  
; NAME/KEY: CDS  
; LOCATION: (4612)...(4737)  
; NAME/KEY: CDS  
; LOCATION: (4827)...(4925)  
; NAME/KEY: CDS  
; LOCATION: (6322)...(6447)  
; NAME/KEY: CDS  
; LOCATION: (6543)...(6638)  
; NAME/KEY: CDS  
; LOCATION: (7137)...(7316)  
; NAME/KEY: CDS  
; LOCATION: (7439)...(7553)  
US-09-167-681-45

Query Match 7.2%; Score 192.8; DB 4; Length 8453;  
Best Local Similarity 75.3%; Pred. No. 5.8e-34;  
Matches 253; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 1525 TTTTGTGTTTATTTTGAATGGAGTCTCGTCTGTGTCGCCAGGCTGAGTGCAGTAG 1584  
Db 2562 TTTTGTGTTTATTTTNGAGTCTCGTCTGTGTCGCCAGGCTGAAGTGCAGTAG 2503  
QY 1585 TGCATCTCGCTCACTACAACTCCACCTCCCTGGGCTCAAGCGATCCCTCCACCTCAG 1644  
Db 2502 CAGCATCTCGCTCACTGTAATCTCCCTCCT-GGTTCAAGCAATCTCTCGCTCAG 2444  
QY 1645 CCGCCAAGTAGCTGGGACTATAGGTGTGTACCATCACACTGGCTAAATTTTGTATTTT 1704  
Db 2443 CTTCTGAAATAGTGGGATTTACAGGACCAACACCAAGCTAAATTTTGTATTTT 2384

QY 1705 TTGTAGACACAGGTTTCCGCTGTCGCCAGGCTGCTTGAATCTCTGAGTCAAGCA 1764  
Db 2383 TAGTAGACACAGGTTTTCGCTGTCGCCAGGCTGCTCAAACTCTGAGTCAAGT 2324  
QY 1765 ACCTGCGGCTCGGCTCCCAAGTACTGGGATTACACGAAAGGACCATGCCAGG 1824  
Db 2323 ATCGCCGCTCGGCTCCCAAGTGTGAGATTACAGGATGAGCCCATGCCGTC 2264  
QY 1825 CTAGATGTGCTTATCCCAATCTTTGGCAGGCTG 1860  
Db 2263 CGACAAGCTTCTCTAATTCACCCAGGCAAGGAGG 2228

## RESULT 9

US-09-741-150-3/c  
; Sequence 3, Application US/09741150  
; Patent No. 6436689  
; GENERAL INFORMATION:  
; APPLICANT: GUGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: CL000968  
; CURRENT APPLICATION NUMBER: US/09/741,150  
; CURRENT FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 112132  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(112132)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-741-150-3

Query Match 7.2%; Score 192.4; DB 4; Length 112132;  
Best Local Similarity 78.4%; Pred. No. 1.6e-33;  
Matches 243; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 1526 TTTGTTTTTATTTTGAATGGAGTCTCGTCTGTGTCGCCAGGCTGAGTGCAGTAGT 1585  
Db 97233 TTTTATTTTATTTTGGACAGGCTGTTGTTCTGTCACCCAGGCTGAATGCAGTGGC 97174  
QY 1586 GCAATCTCGCTCACTACAACTCCACTCCCTGGGCTCAAGCGATCTCCACCTCAGC 1645  
Db 97173 GTGATCTTGCTCATTCGAACCTCC-CTCTCCAGGCTCAAGTATCTCCACCTCAGC 97115  
QY 1646 CGCGAAGTAGTGGGACTATAGGTGTGTACCATCACCTGGCTAATTTTGTATTTT 1705  
Db 97114 CTCCCAAGTAGTTGTACTACAGGCATAAGCCACACACAGCTAATTTTGTATTTT 97055  
QY 1706 TGTAGACACAGGTTTCCGCTGTCGCCAGGCTGTTGAAATCTCTGAGTCAAGCAA 1765  
Db 97054 TGTAGAAGGATTTTGGCTGTCGCCAGGCTGTTGAACTCTGAGTCAAGCAA 96995  
QY 1766 CTTGCGGCTCGGCTCCCAAGTACTGGGATTACACGAAAGGACCATGCCAGG 1825  
Db 96994 TCCACCTGCTCAGCTCCCAAGTGTGGGATTACAGGATGAGCCACCGCCAGC 96935  
QY 1826 TAGATGTCT 1835  
Db 96934 TGGATATTC 96925

## RESULT 10

US-08-816-617A-1  
; Sequence 1, Application US/08816617A  
; Patent No. 6022741  
; GENERAL INFORMATION:  
; APPLICANT: Tling, Jenny P.-Y.  
; APPLICANT: Piskurich, Janet

```

; TITLE OF INVENTION: No. 6022741el Regulatory Genetic DNA that
; TITLE OF INVENTION: Regulates the Class II Transactivator (CIITA)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6022741th Carolina
; COUNTRY: United States
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,617A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6678 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-816-617A-1

Query Match          7.28; Score 192.2; DB 3; Length 6678;
Best Local Similarity 77.88; Pred. No. 7.3e-34;
Matches 245; Conservative 0; Mismatches 68; Indels 2; Gaps 1;

QY 1525 TTTTGTATTTATTTTGAATGGAGTCTCGTTCTCTGCGCCAGGCTGAGGTGACGTAG 1584
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3043 TTTTGTATTTATTTTGAATAGAGTCTCGCTCTGAAGCCAGGCTGGAGTGCAGTGG 3102

QY 1585 TGAATCTCCGCTCACT--ACAACCTCCACTCCCTGGGGCTCAAGCATCTCCACCTC 1642
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3103 TGTGATCTCGGCTCACTTGCACCTCCACCTCCCGAGGTTCAAGCAATCTCTCGCTC 3162

QY 1643 AGCCGCGAAGTAGCTGGGACTAGTGTGTACCATCACACCTGGGCTAATTTTGTATT 1702
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3163 AGCTTCCCAAGTACTGGGATTACAGGGGTGCACACACACCTGGATATTTTCTATT 3222

QY 1703 TTTTGTAGACACAGGGTTTCCCAATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAG 1762
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3223 TTTAGTAGACAGAGGTTTACCATGTTGGCCAGGCTGGTTTCCGATTCCTGACCTCAGG 3282

QY 1763 CAACCTGCCGCTCGCCCTCCCAAGTACTGGGATTACAGGAGAGGACCATGCCCCA 1822
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3283 TGATCCGCGCCGCTCAGTCTCTTAAAGTGTGAGATTACAGGGCTGAGCCACGCGCG 3342

QY 1823 GCCTAGATGTGCTTT 1837
      ||||| |||||
DB 3343 GCCTGACGAGTCTTT 3357

RESULT 11
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.

```

```

; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 9301
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-449-218D-18

Query Match          7.28; Score 192.2; DB 4; Length 9301;
Best Local Similarity 77.48; Pred. No. 8.1e-34;
Matches 233; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1525 TTTTGTATTTATTTTGAATGGAGTCTCGTTCTCTGCGCCAGGCTGAGGTGACGTAG 1584
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6477 TGTGTGTTTGTGTTTGGAGACAGAGTCTCGCTCTATTGCCAGGCTGGAGTGCAGTGT 6536

QY 1585 TGAATCTCCGCTCACTACACCTCCACTCCCTGGGGCTCAAGGATCTCCACCTCAG 1644
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6537 CACAATCTCGGCTTACTGCAACTTCTGCTTCCCGGATTCAAGTGATTCTCTCGCTCAG 6596

QY 1645 CCGCGAAGTAGCTGGGACTATAGTGTGTACCATCACACTGGCTAATTTTGTATTTT 1704
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6597 CTTCCAGAGTAGCTGGGATTACAGGTCGTCGCCACACGCTGGCTAATTTTGTATTTT 6656

QY 1705 TTGTAGACACAGGTTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6657 TGATAGACAGGGGTTTTCACCATGTTGCCAGGCTAGTCTCGAACTCTTGACCTCAAGTG 6716

QY 1765 ACCTGCCGCGCTCGGCCCTCCCAAGTACTGGGATTACAGGAGGACCATGCCCAGG 1824
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6717 ATCTGCTGCTCGGCCCTCCCAAGTCTGGGATTACAGGCTGAGGACCATGCCCAGG 6776

QY 1825 C 1825
DB 6777 C 6777

RESULT 12
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:

```



```
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 36,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 35060 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Cosmid including ACHE
DESCRIPTION: promoter, ACHE gene and ARS gene"
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 7q22
FEATURE:
NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: 34895..34895
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: gene"
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 34895..34895
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 33779..33963
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 33493..33591
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 33297..33408
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: 32959..33094
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: 32386..32468
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: 31894..32080
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: 31363..31534
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: 31131..31284
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: 30816..31011
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: 30470..30626
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: 30187..30274
OTHER INFORMATION: /gene= "AR"
```

OTHER INFORMATION: /number= 14  
FEATURE:  
NAME/KEY: exon  
LOCATION: complement (29945..30073)  
OTHER INFORMATION: /gene= "AR"  
OTHER INFORMATION: /number= 15  
FEATURE:  
NAME/KEY: exon  
LOCATION: complement (29664..29856)  
OTHER INFORMATION: /gene= "ARS"  
OTHER INFORMATION: /number= 16  
US-08-814-095-7

Query Match 7.2%; Score 192.2; DB 3; Length 35060;  
Best Local Similarity 77.4%; Pred. No. 1.3e-33;  
Matches 233; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
QY 1530 TTTTATTTTGTAAATGGAGTCTCGTTCTGTCGCCAGGCTGAGGTGCAGTAGTCA 1589  
DB 14160 TTTTATTTTGTAGCGGAGTTAGCTCTGTCAACCAGGCTGGAGTGCAGTGGCGTGA 14101  
QY 1590 TCTCGGCTCACTACAACCTCCACCTCCCTGGGGCTCAAGCGATCCTCCACCTCAGCCGCC 1649  
DB 14100 TCTCAATCACTGTAACCTCCGCCGCCAGAGTTCAAGCGATTCTCTGCTCAGCCTCC 14041  
QY 1650 GAAGTAGCTGGACTATAGGTGTGTACCATCACCTGGCTAAATTTTGTATTTTGTGA 1709  
DB 14040 CAAGTAGCTGGGATTACAGCGCCCTGCCACACCTGGCTAAATTTTCCATTTTGTGA 13981  
QY 1710 GACACAGGTTTTCGCTGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTG 1769  
DB 13980 GAGCAGGTTTTCGCTGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTG 13921  
QY 1770 CCGGCTCGGCTCCCAAGTACTGGGATTACAGCAAGGACCATGCCAGGCTAGA 1829  
DB 13920 CTTGCTGGCTTCCCAAGTACTGGGATTACAGCAAGGACCATGCCAGGCTAGA 1829  
QY 1830 T 1830  
DB 13860 T 13860

RESULT 13  
US-09-791-211-10  
; Sequence 10, Application US/09791211  
; Patent No. 6448080  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
; FILE REFERENCE: RTS-0205  
; CURRENT APPLICATION NUMBER: US/09/791,211  
; CURRENT FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 90  
; SEQ ID NO 10  
; LENGTH: 9884  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
NAME/KEY: unsure  
LOCATION: 24962  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 64383  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65468  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65469  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65470

OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65471  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 87130  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 89049  
OTHER INFORMATION: unknown  
OTHER INFORMATION:  
US-09-791-211-10

Query Match 7.2%; Score 192.2; DB 4; Length 98844;  
Best Local Similarity 79.0%; Pred. No. 1.7e-33;  
Matches 241; Conservative 0; Mismatches 63; Indels 1; Gaps 1;  
QY 1526 TTTGTTTTTATTTTGTAAATGGAGTCTCGTTCTGTCGCCAGGCTGAGGTGCAGTAGT 1585  
DB 2480 TTTTATTTTGTAGATGGCTCTCACTCCGTTGCCAGGAGTGCAGTGGT 2539  
QY 1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCACCTCAGC 1645  
DB 2540 GCCATCTAGCTCACTGCAACCTCCGCTCCT-AGGTTCAAGCGATTCTCTGCTCAGC 2598  
QY 1646 CGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAAATTTTGTATTTT 1705  
DB 2599 CTTCTGACTAGCTGGGATTACAGGAGGCTGCCACCATCCCTGGCTAAATTTTGTATTTT 2658  
QY 1706 TGTAGACACAGGTTTTCGCTGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAA 1765  
DB 2659 AGTAGAGACGGAGTTTAAACCATGTTGGTCAAGCTGGTCTTGAATTCCTGAGCTCAAGTGA 2718  
QY 1766 CTTCCGCGCTTCGCGCTCCCAAGTACTGGGATTACAGCAAGGACCATGCCAGGC 1825  
DB 2719 TCCGCGCTTCGCGCTCCCAAGTACTGGGATTACAGCAAGGACCATGCCAGGC 2778  
QY 1826 TAGAT 1830  
DB 2779 TAAAT 2783

RESULT 14  
US-09-729-995-3/c  
; Sequence 3, Application US/09729995  
; Patent No. 6426206  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: THEREOF  
; FILE REFERENCE: CL000904  
; CURRENT APPLICATION NUMBER: US/09/729,995  
; CURRENT FILING DATE: 2000-12-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 29629  
; TYPE: DNA  
; ORGANISM: Human  
US-09-729-995-3

Query Match 7.2%; Score 191.4; DB 4; Length 29629;  
Best Local Similarity 76.3%; Pred. No. 1.8e-33;  
Matches 248; Conservative 0; Mismatches 76; Indels 1; Gaps 1;  
QY 1524 GTTTTGTATTTTATTTTGTAAATGGAGTCTCGTTCTGTCGCCAGGCTGAGGTGCAGTA 1583  
DB 11930 GTTTTGTATTTTATTTTGTAAATGGAGTCTCGTTCTGTCGCCAGGCTGAGGTGCAGTA 11871  
QY 1584 GTCAATCTCGGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCACCTCA 1643  
DB 11870 GCGTAGTCTGGCTCACTGCAAGCTCCAC-CTCCAGGTTCAAGAGATTCTCTGCTCCTCA 11812

|    |       |     |  |       |
|----|-------|-----|--|-------|
| Qy | 1644  | GC  | CGCCGAAGTAGCTAGCTGGGACTATAGGTGTGTACCATCACACCTGCTGAATTTTGTGATTT | 1703  |
| Db | 11811 | GC  | CTCCGAGTAGTTGGGATTACAGGTGTCTGCCACACGCCGGCTACTTTTTGTGATTT       | 11752 |
| Qy | 1704  | TT  | TGTAGACACAGGGTTTCGCCATCTTGCCAGGCTGGTCTGTGAATTCCTCAGCTCAAGC     | 1763  |
| Db | 11751 | TT  | AGTAGACACAGGGTTTCGCCATCTTGCTCAGGCTGGTCTGTGAATTCCTCAGCTCAGGT    | 11692 |
| Qy | 1764  | AC  | CTCGCGGCTCGGCTCTCCCAAGTAGTCTGGGATTACACGCAGAGCCACCATGCCCG       | 1823  |
| Db | 11691 | GAT | CCACCCACCTCGGCTCTCCCAAGTAGTCTGGGATTACAGCGGTAGCCACCGCCCCG       | 11632 |
| Qy | 1824  | GCT | AGATGTGCTTATCCCAATCCT  | 1848  |
| Db | 11631 | CT  | GGGAGCAGAGTTCTTCTTAGGCT  | 11607 |

## RESULT 15

US-08-480-784-20  
Sequence 20, Application US/08480784  
Patent No. 5693473  
GENERAL INFORMATION:  
APPLICANT: Skolnick, Mark H.  
APPLICANT: Goldgar, David E.  
APPLICANT: Miki, Yoshio  
APPLICANT: Swenson, Jeff  
APPLICANT: Kamb, Alexander  
APPLICANT: Harshman, Keith D.  
APPLICANT: Shattuck-Eidens, Donna M.  
APPLICANT: Tavtvgian, Sean W.  
APPLICANT: Wiseman, Roger W.  
APPLICANT: Futreal, P. Andrew  
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer  
Susceptibility Gene  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,784  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/409,305  
FILING DATE: 24-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/348,824  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/308,104  
FILING DATE: 16-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/300,266  
FILING DATE: 02-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,221  
FILING DATE: 12-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109347  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810

```

; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6769 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
US-08-480-784-20

Query Match 7.1%; Score 190.8; DB 1; Length 6769;
Best Local Similarity 75.5%; Pred. NO. 1.5e-33;
Matches 237; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1526 TTTGTTTTTTATTTTTTGAATGGAGTCCTCGTTCGTCCGCCAGGCTGAGGTGCAGTACT 1585
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6242 TTTGTTTGTGTTTTTTTGATGGGGTCTCACTCTGTTGCCAGGCTGGAGTGCAGTGAT 6301
QY 1586 GCAATCTCGCTCACTACAACTCCACTCCCTGGGGCTCAAGGCGATCCTCCCACTCAGC 1645
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6302 GCAATCTTGCTGCACATGCAACCTCCACCTCCAAAGGCTCAAGGTATCCTCCCACTTCAGC 6361
QY 1646 CGCCGAAGTAGCTGGGAGCTATAGGTGTGTAACCATCACACTGGCTAATTTTTGTATTTTT 1705
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6362 CTCCTGAGTAGCTGGGAGCTACAGGCGCATGCCACCACACCCGGTTAATTTTTGTGGTTT 6421
QY 1706 TGTGACACACAGGGTTTTCGCCATGTTGCCCAGGCTGCTGTGAATTCCTGAGCTCAAGCAA 1765
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6422 TATGAGATGGGGTTTTCACCATGTTTACCAGGGTGTGCTCAAACTCCTGGACTCAAGCAG 6481
QY 1766 CTTGCCGGCTCGGGCTCCCAAAGTACTCGGAGTTACACGAAGGCACCATGCCCGAGGC 1825
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6482 TCTGCCCACTTCAGCCTCCCAAAGTGCATGAGGTTACAGGCTGAGCCATGTGCCCTGGCC 6541
QY 1826 TAGAATGTCATTAT 1839
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
6542 TGCCCTTTACTTTT 6555
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: June 24, 2003, 00:53:00  
Job time : 124 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:56:39 ; Search time 376 Seconds  
(without alignments)  
10435.905 Million cell updates/sec

Title: us-09-817-199b-1

Perfect score: 2674  
Sequence: 1 ttcgctggcgccggcact.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length  | ID | Description         |
|------------|--------|-------------|---------|----|---------------------|
| 1          | 2674   | 100.0       | 2674    | 10 | US-09-817-199A-1    |
| 2          | 2581.6 | 96.5        | 2623    | 9  | US-09-764-868-71    |
| 3          | 2031.8 | 78.0        | 13182   | 10 | US-09-817-199A-3    |
| 4          | 1286   | 48.1        | 1316    | 9  | US-09-764-868-493   |
| 5          | 1099.6 | 41.1        | 1106    | 9  | US-10-037-270-959   |
| 6          | 1086   | 40.6        | 1116    | 10 | US-09-794-257-13    |
| 7          | 873    | 32.6        | 875     | 12 | US-10-051-986-10    |
| 8          | 683.4  | 25.6        | 24707   | 9  | US-10-274-968-3     |
| 9          | 683.4  | 25.6        | 24707   | 10 | US-09-740-027-3     |
| 10         | 600.6  | 22.5        | 601     | 10 | US-09-817-199A-24   |
| 11         | 576    | 21.5        | 576     | 10 | US-09-794-257-15    |
| 12         | 573    | 21.4        | 601     | 10 | US-09-817-199A-25   |
| 13         | 443.4  | 16.6        | 447     | 10 | US-09-867-550-1811  |
| 14         | 415    | 15.5        | 417     | 9  | US-09-918-995-34999 |
| 15         | 233.4  | 8.7         | 434     | 9  | US-09-918-995-648   |
| 16         | 209.4  | 7.8         | 964     | 9  | US-09-764-868-75    |
| 17         | 204.8  | 7.7         | 10859   | 9  | US-09-764-891-7408  |
| 18         | 203.6  | 7.6         | 36221   | 9  | US-09-954-556-29    |
| 19         | 200.8  | 7.5         | 1691139 | 9  | US-10-067-514-1     |

c

|      |       |     |        |    |                     |                   |
|------|-------|-----|--------|----|---------------------|-------------------|
| c 20 | 200.6 | 7.5 | 21833  | 10 | US-09-764-877-2275  | Sequence 2275, Ap |
| c 21 | 199.8 | 7.5 | 1130   | 10 | US-09-764-877-3373  | Sequence 3373, Ap |
| c 22 | 198.6 | 7.4 | 15745  | 10 | US-09-764-877-3236  | Sequence 3236, Ap |
| c 23 | 198.6 | 7.4 | 17216  | 10 | US-09-764-877-3565  | Sequence 3565, Ap |
| c 24 | 198.6 | 7.4 | 17217  | 10 | US-09-764-877-3566  | Sequence 3566, Ap |
| c 25 | 198.4 | 7.4 | 4962   | 9  | US-10-091-504-1902  | Sequence 1902, Ap |
| c 26 | 198.4 | 7.4 | 4962   | 9  | US-10-091-504-1903  | Sequence 1903, Ap |
| c 27 | 198.4 | 7.4 | 4962   | 9  | US-09-764-869-1902  | Sequence 1902, Ap |
| c 28 | 198.4 | 7.4 | 4962   | 9  | US-09-764-869-1903  | Sequence 1903, Ap |
| c 29 | 198   | 7.4 | 32205  | 9  | US-09-764-891-10213 | Sequence 10213, A |
| c 30 | 198   | 7.4 | 32205  | 9  | US-10-205-428-1012  | Sequence 1012, Ap |
| c 31 | 197.6 | 7.4 | 7791   | 10 | US-09-764-877-3700  | Sequence 3700, Ap |
| c 32 | 197.4 | 7.4 | 65484  | 9  | US-09-859-888-3     | Sequence 3, Appli |
| c 33 | 197.2 | 7.4 | 2754   | 10 | US-09-745-763-133   | Sequence 133, App |
| c 34 | 197.2 | 7.4 | 15515  | 10 | US-09-822-860-3     | Sequence 3, Appli |
| c 35 | 196.8 | 7.4 | 17335  | 9  | US-10-092-154-1280  | Sequence 1280, Ap |
| c 36 | 196.8 | 7.4 | 17335  | 10 | US-09-764-847-1280  | Sequence 1280, Ap |
| c 37 | 196.8 | 7.4 | 19882  | 9  | US-10-092-154-1281  | Sequence 1281, Ap |
| c 38 | 196.8 | 7.4 | 19882  | 10 | US-09-764-847-1281  | Sequence 1281, Ap |
| c 39 | 196.6 | 7.4 | 13821  | 10 | US-09-764-877-2595  | Sequence 2595, Ap |
| c 40 | 196.2 | 7.3 | 7960   | 9  | US-10-091-504-2327  | Sequence 2327, Ap |
| c 41 | 196.2 | 7.3 | 7960   | 10 | US-09-764-869-2327  | Sequence 2327, Ap |
| c 42 | 196.2 | 7.3 | 19167  | 9  | US-09-764-891-8028  | Sequence 8028, Ap |
| c 43 | 195.8 | 7.3 | 11960  | 9  | US-09-764-891-8501  | Sequence 8501, Ap |
| c 44 | 195.8 | 7.3 | 203654 | 10 | US-09-820-905-3     | Sequence 3, Appli |
| c 45 | 195.4 | 7.3 | 30175  | 10 | US-09-738-878-3     | Sequence 3, Appli |

#### ALIGNMENTS

##### RESULT 1

US-09-817-199A-1

; Sequence 1, Application US/09817199A

; Patent No. US20020142380A1

; GENERAL INFORMATION:

; APPLICANT: SHAC, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01187

; CURRENT APPLICATION NUMBER: US/09/817,199A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2674

; TYPE: DNA

; ORGANISM: Human

US-09-817-199A-1

|                       |              |  |            |              |
|-----------------------|--------------|--|------------|--------------|
| Query Match           | 100.0%;      | Score 2674;  | DB 10;     | Length 2674; |
| Best Local Similarity | 100.0%;      | Pred. No. 0;   |            |              |
| Matches 2674;         | Conservative | 0;   | Mismatches | 0;           |
|                       |              |  | Indels     | 0;           |
|                       |              |  | Gaps       | 0;           |
| Qy                    | 1            | TTCCCTTCGCGCGCGGACATGCTCCTCTCGTCCAGGACATGACGGGACGCCAGCGG       | 60         |              |
| Db                    | 1            | TTCCCTTCGCGCGCGGACATGCTCCTCTCGTCCAGGACATGACGGGACGCCAGCGG       | 60         |              |
| Qy                    | 61           | CCGTTCGCACCCCGGATGGCGGCGCCCGCCCGGCGCTCCCGCCCTGCAGTCCGAGCTACG   | 120        |              |
| Db                    | 61           | CCGTTCGCACCCCGGATGGCGGCGCCCGCCCGGCGCTCCCGCCCTGCAGTCCGAGCTACG   | 120        |              |
| Qy                    | 121          | ACCTTCAGGGCAAGGTGATGCTTCTTGGGAGACACAGCGCTCGGCAAAACATGTTTCTCTGA | 180        |              |
| Db                    | 121          | ACCTTCAGGGCAAGGTGATGCTTCTTGGGAGACACAGCGCTCGGCAAAACATGTTTCTCTGA | 180        |              |
| Qy                    | 181          | TCCAATTCAAAGACGGGGCCCTTCTGTCCGAACTTCATAGCCACCGTCCGCATAGACT     | 240        |              |
| Db                    | 181          | TCCAATTCAAAGACGGGGCCCTTCTGTCCGAACTTCATAGCCACCGTCCGCATAGACT     | 240        |              |
| Qy                    | 241          | TCAGGAACAGGTGGTGTGACTGTGGATGGCGCTGAGAGTGAAGCTGCACATCTGGGACACCG | 300        |              |
|                       |              |  |            |              |

Db 241 TCAGGAACAAGGTGGTGAAGTGGATGGCTGAGAGTGAAGCTGCAGATCTGGGACACCG 300  
QY 301 CTGGCAGGAACGGTTCCGAAGCGTACCCAGTGTATTACAGAGATGCTCAGGCGCTTGC 360  
Db 301 CTGGCAGGAACGGTTCCGAAGCGTACCCAGTGTATTACAGAGATGCTCAGGCGCTTGC 360  
QY 361 TTCGTGCTATGACATCACCAACAAATCTCTTTTCGACAAATCAGGCGCTGGCTCACTG 420  
Db 361 TTCGTGCTATGACATCACCAACAAATCTCTTTTCGACAAATCAGGCGCTGGCTCACTG 420  
QY 421 AGATTATGATATGCCAGAGGAGCGTGGTGATGCTGCTAGGCAACAAAGGGGATA 480  
Db 421 AGATTATGATATGCCAGAGGAGCGTGGTGATGCTGCTAGGCAACAAAGGGGATA 480  
QY 481 TGAGCAGGAAAGAGTATCGTTCCGAGAGGAGAGACCTTGGCCAGGAGTACGGTG 540  
Db 481 TGAGCAGGAAAGAGTATCGTTCCGAGAGGAGAGACCTTGGCCAGGAGTACGGTG 540  
QY 541 TTCCTTCTCTGGAGACAGCCCAAGACTGGCATGAATGTGGAGTTAGCCCTTCTGGCCA 600  
Db 541 TTCCTTCTCTGGAGACAGCCCAAGACTGGCATGAATGTGGAGTTAGCCCTTCTGGCCA 600  
QY 601 TCGCCAGGAACTGAATATACCGGCGGGGATCATAGGCGGATGAGCCAGCTTCCAGATCC 660  
Db 601 TCGCCAGGAACTGAATATACCGGCGGGGATCATAGGCGGATGAGCCAGCTTCCAGATCC 660  
QY 661 GAGACTATGATAGTATCCAGAGAGCGCTCCAGCTGCTGCTCTCATGTGAATCCAG 720  
Db 661 GAGACTATGATAGTATCCAGAGAGCGCTCCAGCTGCTGCTCTCATGTGAATCCAG 720  
QY 721 GGGCAGAGAGAGGCTCTGGAGGCACACAGGATGAGCCCTTCCAGGCGCTGGC 780  
Db 721 GGGCAGAGAGAGGCTCTGGAGGCACACAGGATGAGCCCTTCCAGGCGCTGGC 780  
QY 781 TTATTTCAAGAGGTGAGCCAAATGGGAGAAAGATGAGGACTCACTGCACAGCGCTTC 840  
Db 781 TTATTTCAAGAGGTGAGCCAAATGGGAGAAAGATGAGGACTCACTGCACAGCGCTTC 840  
QY 841 CTAGCAGGAGCTATCTCAACCTCTACTTGTAGTCTCGGTCTCCCGCATCCACAG 900  
Db 841 CTAGCAGGAGCTATCTCAACCTCTACTTGTAGTCTCGGTCTCCCGCATCCACAG 900  
QY 901 GGAGGTAAACACTTACCTTTTATTTTAAATAGTACATAATTTAAATACCAAAAAGCGC 960  
Db 901 GGAGGTAAACACTTACCTTTTATTTTAAATAGTACATAATTTAAATACCAAAAAGCGC 960  
QY 961 CTGATCCCAAAAACCGAGGTGGAGCTAGTGGCCCTTTGCTTTCTAGGACTTGG 1020  
Db 961 CTGATCCCAAAAACCGAGGTGGAGCTAGTGGCCCTTTGCTTTCTAGGACTTGG 1020  
QY 1021 GGGCGGCGCTTCCCTCTTAAGCATAAACAAAGGTGGTGTGCTCCAGCTCAGCCCGAGGG 1080  
Db 1021 GGGCGGCGCTTCCCTCTTAAGCATAAACAAAGGTGGTGTGCTCCAGCTCAGCCCGAGGG 1080  
QY 1081 ACACAGATGACATTTGGGGTGAGGCGAGTAATGACTCCATCGACCCCTCAGTTCAGCT 1140  
Db 1081 ACACAGATGACATTTGGGGTGAGGCGAGTAATGACTCCATCGACCCCTCAGTTCAGCT 1140  
QY 1141 GGACAGAGGCTCAGGTACCCCGCTTCACTGTCTCCCGCTTCCAGGAGCTTATCTTC 1200  
Db 1141 GGACAGAGGCTCAGGTACCCCGCTTCACTGTCTCCCGCTTCCAGGAGCTTATCTTC 1200  
QY 1201 GCCCATCTCCCAATTAAGTGGCGCTTGTGCTGTAGGAGAACCAAGCGCTCAGGGAAG 1260  
Db 1201 GCCCATCTCCCAATTAAGTGGCGCTTGTGCTGTAGGAGAACCAAGCGCTCAGGGAAG 1260  
QY 1261 ATAAGAGATATGGAGATGGAGGGGAGGACAAAGGGGACAGAGTAGGGTCTAGCTGGCT 1320  
Db 1261 ATAAGAGATATGGAGATGGAGGGGAGGACAAAGGGGACAGAGTAGGGTCTAGCTGGCT 1320  
QY 1321 ATCTCTGGCCTTACTACACCCCTTGGAGGATGCGCCCTTTCTCCAGCACACAAGCAC 1380  
Db 1321 ATCTCTGGCCTTACTACACCCCTTGGAGGATGCGCCCTTTCTCCAGCACACAAGCAC 1380

QY 1381 ATTGGGCACTGGAATAATTGGTTCCAGGCTCCTGTCTCTGGACTTCAGATCCTGGGG 1440  
Db 1381 ATTGGGCACTGGAATAATTGGTTCCAGGCTCCTGTCTCTGGACTTCAGATCCTGGGG 1440  
QY 1441 GAGCCCTTCCCGCCCTGAATCCCTGGCTTAGCTACCTTCCCTGCTGTGACCTTAAAC 1500  
Db 1441 GAGCCCTTCCCGCCCTGAATCCCTGGCTTAGCTACCTTCCCTGCTGTGACCTTAAAC 1500  
QY 1501 CTCAGGTCAGAACTAGGAAAGAGTGTGTTTTATTTTATTTTGAATGGAGTCTCGTCT 1560  
Db 1501 CTCAGGTCAGAACTAGGAAAGAGTGTGTTTTATTTTATTTTGAATGGAGTCTCGTCT 1560  
QY 1561 GTCCGCCAGGCTGAGTGCAGTAGTCAATCTCCGCTCAGCTACCACTTCCACTCCCTGGG 1620  
Db 1561 GTCCGCCAGGCTGAGTGCAGTAGTCAATCTCCGCTCAGCTACCACTTCCACTCCCTGGG 1620  
QY 1621 GCTCAAGGATCCCTCCAGCTCAGCCGCGGAGTAGCTGGGACTATAGGTGTGTACCATC 1680  
Db 1621 GCTCAAGGATCCCTCCAGCTCAGCCGCGGAGTAGCTGGGACTATAGGTGTGTACCATC 1680  
QY 1681 ACACCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTTCGCTATGTTGCCAGGCTG 1740  
Db 1681 ACACCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTTCGCTATGTTGCCAGGCTG 1740  
QY 1741 GTCTTGAATTCCTGAGCTCAAGCAACTCGCGGCTCGGCTCCCAAGTACTGGGATTA 1800  
Db 1741 GTCTTGAATTCCTGAGCTCAAGCAACTCGCGGCTCGGCTCCCAAGTACTGGGATTA 1800  
QY 1801 CACSCAAGGACCACTGCCCAGCTAGATGCTCTTATCCCATCTTTGGGAGGATG 1860  
Db 1801 CACSCAAGGACCACTGCCCAGCTAGATGCTCTTATCCCATCTTTGGGAGGATG 1860  
QY 1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTGAGCCCTCCTGGTTAAGAGCC 1920  
Db 1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTGAGCCCTCCTGGTTAAGAGCC 1920  
QY 1921 AGATAAGGAAATCCCTTTCTAGTGTGGAATGTGTTGAAAAAAGAGAAATCCC 1980  
Db 1921 AGATAAGGAAATCCCTTTCTAGTGTGGAATGTGTTGAAAAAAGAGAAATCCC 1980  
QY 1981 TGGCTCTCGAGCTGGTGGGAGCAAGATTAAAGCAACCTCCCTGACATGTATCCCTTT 2040  
Db 1981 TGGCTCTCGAGCTGGTGGGAGCAAGATTAAAGCAACCTCCCTGACATGTATCCCTTT 2040  
QY 2041 GACCCCAAGCTTCCCTTCCCTGACCCCATGCGCTTTCTTAACTTCTCAACAG 2100  
Db 2041 GACCCCAAGCTTCCCTTCCCTGACCCCATGCGCTTTCTTAACTTCTCAACAG 2100  
QY 2101 ATACCCAGGCTTAACCTGCTTTACCTCCCTCTACTAGTCAAGTGTAGTGGGAGG 2160  
Db 2101 ATACCCAGGCTTAACCTGCTTTACCTCCCTCTACTAGTCAAGTGTAGTGGGAGG 2160  
QY 2161 TCACCATTTTCCGAGTTAAACCAATGCAATATGAGTAAAAAAGTCACTGTGGTATGTC 2220  
Db 2161 TCACCATTTTCCGAGTTAAACCAATGCAATATGAGTAAAAAAGTCACTGTGGTATGTC 2220  
QY 2221 TGGGGTAGAGAGGGGTAGCAATGCTGCTCTCTTGGTACATATCTCCCAAGC 2280  
Db 2221 TGGGGTAGAGAGGGGTAGCAATGCTGCTCTCTTGGTACATATCTCCCAAGC 2280  
QY 2281 TCTGATCCCTGCGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340  
Db 2281 TCTGATCCCTGCGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340  
QY 2341 ACTGCAAGCTTCCGCGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400  
Db 2341 ACTGCAAGCTTCCGCGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400  
QY 2401 CTGAGGCTACTGATTTGCTTTCAAGGCAAAATCTTGTCTCAGCAGTCAAGGCTTCCA 2460  
Db 2401 CTGAGGCTACTGATTTGCTTTCAAGGCAAAATCTTGTCTCAGCAGTCAAGGCTTCCA 2460

[illegible]

## RESULT 2

US-09-764-868-71

```

: Sequence 71, Application US/09764868
: Patent No. US20020168711A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT32
: CURRENT APPLICATION NUMBER: US/09/764,868
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1510
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 71
: LENGTH: 2623
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-764-868-71

```

| Query Match           | 96.5%        | Score 2581.6  | DB 9       | Length 2623 |
|-----------------------|--------------|---|------------|-------------|
| Best Local Similarity | 99.8%        | Pred. No. 0   |            |             |
| Matches 2606          | Conservative | 0   | Mismatches | 4           |
|                       |              |   | Indels     | 2           |
|                       |              |   | Gaps       | 2           |
| Qy                    | 42           | ATGACGGGCACGCAGGCGCGTGTGCACCCGGATGGCGAGGACCCCGAGCGCTCCCCG     | 101        |             |
| Db                    | 12           | ATGACGGGCACGCAGGCGCGTGTGCACCCGGATGGCGAGGACCCCGAGCGCTCCCCG     | 71         |             |
| Qy                    | 102          | CCCTGCAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC   | 161        |             |
| Db                    | 72           | CCCTGCAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC   | 131        |             |
| Qy                    | 162          | GGCAAAACATGTTTCTTGATCCAAATTCAAAGACGGGGCCTTCTGTCCGGAACCTTCATA  | 221        |             |
| Db                    | 132          | GGCAAAACATGTTTCTTGATCCAAATTCAAAGACGGGGCCTTCTGTCCGGAACCTTCATA  | 191        |             |
| Qy                    | 222          | GCACCGTCGGCATAGACTTCAGGAAACAGGTGGTGACTGTGTGATGCGGTGAGAGTGAAG  | 281        |             |
| Db                    | 192          | GCACCGTCGGCATAGACTTCAGCAACAAAGTGGTGACTGTGTGATGCGGTGAGAGTGAAG  | 251        |             |
| Qy                    | 282          | CTCAGATCTGGGACACCGCTGGGCGAGAACGGTTCGGAAGCGCTCACCCATGCTATTAC   | 341        |             |
| Db                    | 252          | CTCAGATCTGGGACACCGCTGGGCGAGAACGGTTCGGAAGCGCTCACCCATGCTATTAC   | 311        |             |
| Qy                    | 342          | AGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATCACCACAAATATCTTCTTCGACAAC | 401        |             |
| Db                    | 312          | AGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATCACCACAAATATCTTCTTCGACAAC | 371        |             |
| Qy                    | 402          | ATCAGGCGCTGCTCACTGAGATTCATGAGTATGCCAGAGGACGCTGCTGATCATGCTG    | 461        |             |
| Db                    | 372          | ATCAGGCGCTGCTCACTGAGATTCATGAGTATGCCAGAGGACGCTGCTGATCATGCTG    | 431        |             |
| Qy                    | 462          | CTAGGCACAAAGCGGGATATGACGACGGAAGAGTGTATCCGTTCCGAAGACGGAGAGACC  | 521        |             |
| Db                    | 432          | CTAGGCACAAAGCGGGATATGACGACGGAAGAGTGTATCCGTTCCGAAGACGGAGAGACC  | 491        |             |
| Qy                    | 522          | TTGGCCACGGGAGTACGGTGTTCCTTCCTGGAGACACGCGCCCAAGACTGGCATGAATGTG | 581        |             |

Db 1572 TACAACTCCACTCCCTGGGCTCAAGGATCCCTCCCACTCAGCCGCCCGGAGTAGCTGG 1631  
QY 1661 GACTATAGGTGTACCAATCACACTGGCTAATTTTGTATTTTGTAGACACAGGTT 1720  
Db 1632 GACTATAGGTGTACCAATCACACTGGCTAATTTTGTATTTTGTAGACACAGGTT 1691  
QY 1721 TCGCATCTTCCAGGCTGCTGTGAATCTCTAGCTCAAGCAACCTCGCGGCTGGC 1780  
Db 1692 TCGCATCTTCCAGGCTGCTGTGAATCTCTAGCTCAAGCAACCTCGCGGCTGGC 1751  
QY 1781 CTCCCAAGTACTGGGATTAACACAGAGGACCATGCCAGGCTAGATGTCTTATC 1840  
Db 1752 CTCCCAAGTACTGGGATTAACACAGAGGACCATGCCAGGCTAGATGTCTTATC 1811  
QY 1841 CCAATCTTTGGAGGATGAGCTCCACAGGCGATTCTTCAAGAGCTGAAGTGTTA 1900  
Db 1812 CCAATCTTTGGAGGATGAGCTCCACAGGCGATTCTTCAAGAGCTGAAGTGTTA 1871  
QY 1901 GCCCTCTGGGTTAAGACCAAGATAAGAGAAATCCCTTTCCTAGGTTTGAATGTGTG 1960  
Db 1872 GCCCTCTGGGTTAAGACCAAGATAAGAGAAATCCCTTTCCTAGGTTTGAATGTGTG 1931  
QY 1961 TG-AAAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGAGACAAAGATTAAAGCAAC 2019  
Db 1932 TGAATAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGAGACAAAGATTAAAGCAAC 1991  
QY 2020 TCCCTCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCTCTCTGACACCCAGCCCT 2079  
Db 1992 TCCCTCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCTCTCTGACACCCAGCCCT 2051  
QY 2080 TTCCTTTAACTTCTCAACAGATACCAAGGCGCTAAACTGTCTTACCTCCCTCTCTACTGA 2139  
Db 2052 TTCCTTTAACTTCTCAACAGATACCAAGGCGCTAAACTGTCTTACCTCCCTCTCTACTGA 2111  
QY 2140 GTCAGGTTAGTGTGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATAGTAA 2199  
Db 2112 GTCAGGTTAGTGTGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATAGTAA 2171  
QY 2200 ACAAGTCATGTGGGTATGTCTGGGTAGAGAGGGGTAGCAAGTTCATGTCTCTCTCT 2259  
Db 2172 ACAAGTCATGTGGGTATGTCTGGGTAGAGAGGGGTAGCAAGTTCATGTCTCTCTCT 2231  
QY 2260 TGCTACATATCTCCCAAGCTCTGATCTCTGCTGCAATGGAGTGGAGGAGGAGGAGG 2319  
Db 2232 TGCTACATATCTCCCAAGCTCTGATCTCTGCTGCAATGGAGTGGAGGAGGAGGAGG 2291  
QY 2320 TCATGACCTCAGGCACTTTACTGCACTCTGCGGCTTGGAGGGGAGGAGGAGGAGG 2379  
Db 2292 TCATGACCTCAGGCACTTTACTGCACTCTGCGGCTTGGAGGGGAGGAGGAGGAGG 2351  
QY 2380 AGAAGTATGCTGCACTTTCTGAGGCTACTGCTATTTCTTCAAGGCAAGAAATCTTC 2439  
Db 2352 AGAAGTATGCTGCACTTTCTGAGGCTACTGCTATTTCTTCAAGGCAAGAAATCTTC 2411  
QY 2440 TCTGAGCAGTCAGGCTCCAGTTTGGGCCCCGATAGGAAGTTCTCGTGGCCCTCCCTCA 2499  
Db 2412 TCTGAGCAGTCAGGCTCCAGTTTGGGCCCCGATAGGAAGTTCTCGTGGCCCTCCCTCA 2471  
QY 2500 GGCAGAGCAGGAGGAGGCTGACATTCAGGCTCTCTTGGGCCCCAAGGAGGTTGCGAG 2559  
Db 2472 GGCAGAGCAGGAGGAGGCTGACATTCAGGCTCTCTTGGGCCCCAAGGAGGTTGCGAG 2531  
QY 2560 GAGATCCATPCCATAGACAGCTCTCTGGGCTCTTGCATTTTGTATTTTTCAGAAATTA 2619  
Db 2532 GAGATCCATPCCATAGACAGCTCTCTGGGCTCTTGCATTTTGTATTTTTCAGAAATTA 2591  
QY 2620 GCAGTATTTGGAAAGCAAAAAA 2651  
Db 2592 GCAGTATTTGGAAAGCAAAAAA 2623

RESULT 3

US-09-817-199A-3

; Sequence 3, Application US/09817199A  
; Patent No. US20020142380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; FILE REFERENCE: C0001187  
; CURRENT APPLICATION NUMBER: US/09/817,199A  
; CURRNT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 13182  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(13182)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-817-199A-3

Query Match 76.08; Score 2031.8; DB 10; Length 13182;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2039; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 589 CCTTCTGGGCATCGCCCAAGAACTGAAATACCGGGCGGCATCAGCGGATGAGCCCA 648  
Db 10060 CCATTGTCTTCTTCAAGGAACTGAAATACCGGGCGGCATCAGCGGATGAGCCCA 10119  
QY 649 GCTTCCAGATCCAGACTATGTAGTCCAGAGAACGCTCCAGTGTCTGCTCTTCA 708  
Db 10120 GCTTCCAGATCCAGACTATGTAGTCCAGAGAACGCTCCAGTGTCTGCTCTTCA 10179  
QY 709 TGTGAATCCAGGGCGAGAGAGGCTCTGAGGACACACAGATGACGCTTCCCTT 768  
Db 10180 TGTGAATCCAGGGCGAGAGAGGCTCTGAGGACACACAGATGACGCTTCCCTT 10239  
QY 769 CCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGGAGGACTCACTG 828  
Db 10240 CCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGGAGGACTCACTG 10299  
QY 829 CACAGCGCTTCTTACGAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGGGTCTC 888  
Db 10300 CACAGCGCTTCTTACGAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGGGTCTC 10359  
QY 889 CCGCATCCACAGGAGGTTAAACACTTACTTTTATTAATAGTACATAATTAATAC 948  
Db 10360 CCGCATCCACAGGAGGTTAAACACTTACTTTTATTAATAGTACATAATTAATAC 10419  
QY 949 CAAAAAGGCGCTGGATCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTT 1008  
Db 10420 CAAAAAGGCGCTGGATCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTT 10479  
QY 1009 CTAGGACTTGGGGGGCGGCGCTTCCCTCTTACGATAACAAAGGTGGTGTCTCAGCT 1068  
Db 10480 CTAGGACTTGGGGGGCGGCGCTTCCCTCTTACGATAACAAAGGTGGTGTCTCAGCT 10539  
QY 1069 CAGCCCCAGGGACACAGATGCACTTTGGGGGTGAGGCGAGGTAATGACTCCATCCGACC 1128  
Db 10540 CAGCCCCAGGGACACAGATGCACTTTGGGGGTGAGGCGAGGTAATGACTCCATCCGACC 10599  
QY 1129 CTCAGTTCAGTGGACAGAGGCTCAGTACCCAGCCCTTCTACTGTCTCCGCTCTCCAG 1188  
Db 10600 CTCAGTTCAGTGGACAGAGGCTCAGTACCCAGCCCTTCTACTGTCTCCGCTCTCCAG 10659  
QY 1189 GAGCTTATCTTCCGCCCATCTCCCAATTAAGTGGGCGCTTGTCTGTAGGAGAGCAAA 1248  
Db 10660 GAGCTTATCTTCCGCCCATCTCCCAATTAAGTGGGCGCTTGTCTGTAGGAGAGCAAA 10719  
QY 1249 GCTCAGGAGATAGAGATATGGAGTGGAGGGGGAGGACAAAGGGCAGAGTAGG 1308  
Db 10720 GCTCAGGAGATAGAGATATGGAGTGGAGGGGGAGGACAAAGGGCAGAGTAGG 10779



```
QY 1309 GTCTAGTGGCTATCTCTGGCTTACTTAACACCCCTGGAGCATGCCCCCTTCTCCCA 1368
Db 10780 GTCTAGTGGCTATCTCTGGCTTACTTAACACCCCTGGAGCATGCCCCCTTCTCCCA 10839
QY 1369 GCACACAAGCACATTTGGGACACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 1428
Db 10840 GCACACAAGCACATTTGGGACACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 10899
QY 1429 CAGATCTGGGGAGGCCCCCTCCGCCCTGAAATCCCTGGCTTAGCTACCTTCTGCTGT 1488
Db 10900 CAGATCTGGGGAGGCCCCCTCCGCCCTGAAATCCCTGGCTTAGCTACCTTCTGCTGT 10959
QY 1489 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAGAGTTTGTATTTTGAATG 1548
Db 10960 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAGAGTTTGTATTTTGAATG 11019
QY 1549 GAGTCGTGTTCTCTGCCCCAGGCTGAGGTGAGTGTGCAATCTCCGCTCAGTACACCT 1608
Db 11020 GAGTCGTGTTCTCTGCCCCAGGCTGAGGTGAGTGTGCAATCTCCGCTCAGTACACCT 11079
QY 1609 CCACCTCCCTGGGCTCAAGGATCTCCACCTCAGCCGCCGAGTAGCTGGGACTATAG 1668
Db 11080 CCACCTCCCTGGGCTCAAGGATCTCCACCTCAGCCGCCGAGTAGCTGGGACTATAG 11139
QY 1669 GTGTGTACCATCACACCTGGCTTAATTTTCTATTTTGTAGACACAGGTTTCGCCATG 1728
Db 11140 GTGTGTACCATCACACCTGGCTTAATTTTCTATTTTGTAGACACAGGTTTCGCCATG 11199
QY 1729 TTGCCAGGCTGTTTGAATCTCTGAGCTCAAGCAACCTGCCGGCTCGGCCCTCCCAA 1788
Db 11200 TTGCCAGGCTGTTTGAATCTCTGAGCTCAAGCAACCTGCCGGCTCGGCCCTCCCAA 11259
QY 1789 GTACTGGGATACACGAGGACGACCATGCCAGGCTAGATGTGCTTATCCCAATCCT 1848
Db 11260 GTACTGGGATACACGAGGACGACCATGCCAGGCTAGATGTGCTTATCCCAATCCT 11319
QY 1849 TTGGCAGGATGAGCTCCACAGGCGATTTCTCAAGCAGCTGAAGTGTATAGCCCTCT 1908
Db 11320 TTGGCAGGATGAGCTCCACAGGCGATTTCTCAAGCAGCTGAAGTGTATAGCCCTCT 11379
QY 1909 GGTTTAAGAGCCAGATAAGGAGAAATCCCTTCTAGTTTGAATGTGTGTGAAGAAA 1968
Db 11380 GGTTTAAGAGCCAGATAAGGAGAAATCCCTTCTAGTTTGAATGTGTGTGAAGAAA 11439
QY 1969 AAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAACTCCCTCTGAC 2028
Db 11440 AAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAACTCCCTCTGAC 11499
QY 2029 ATGTATCCCTTTGACCCCAAGCTCTGCTCTCCCTGACCAACCATGCCCTTCTCTTAA 2088
Db 11500 ATGTATCCCTTTGACCCCAAGCTCTGCTCTCCCTGACCAACCATGCCCTTCTCTTAA 11559
QY 2089 CTTCTCAACAGATACACAGGCTTAACTGCTTTTACCTCCCTCTCTACTGAGTCAGTTA 2148
Db 11560 CTTCTCAACAGATACACAGGCTTAACTGCTTTTACCTCCCTCTCTACTGAGTCAGTTA 11619
QY 2149 GGTGTGGGAGGTCACCATTTCCGAGTTAAACCAATGCAATATAGTAAACAAAGTCA 2208
Db 11620 GGTGTGGGAGGTCACCATTTCCGAGTTAAACCAATGCAATATAGTAAACAAAGTCA 11679
QY 2209 TGTGGGTATGCTCTGGGTAGAGAGAGGGTATGCAAGTTTCTCTCTCTCTGTCACAT 2268
Db 11680 TGTGGGTATGCTCTGGGTAGAGAGAGGGTATGCAAGTTTCTCTCTCTCTGTCACAT 11739
QY 2269 ATCTCCAAAGCTCTGATCCCTGCAATGGGAAGTGGACAGGAACATGAGTCAATGACCT 2328
Db 11740 ATCTCCAAAGCTCTGATCCCTGCAATGGGAAGTGGACAGGAACATGAGTCAATGACCT 11799
QY 2329 GCAGGCATCTTTACTGAGCTCTGCCGGCTTGGAGGGGAGAGGGGAGGAAGATG 2388
Db 11800 GCAGGCATCTTTACTGAGCTCTGCCGGCTTGGAGGGGAGAGGGGAGGAAGATG 11859
```

```
QY 2389 CGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAG 2448
Db 11860 CGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAG 11919
QY 2449 TCAAGCGGCTCCAGTTTGGGCCCCGATAGGAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCA 2508
Db 11920 TCAAGCGGCTCCAGTTTGGGCCCCGATAGGAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCA 11979
QY 2509 GGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCAGGAGATCCAA 2568
Db 11980 GGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCAGGAGATCCAA 12039
QY 2569 TCCCATAGACAGCTCTGGGCTCTTGCAATTTGACTTTTTCAGAAATTAACCTGCAGTATTT 2628
Db 12040 TCCCATAGACAGCTCTGGGCTCTTGCAATTTGACTTTTTCAGAAATTAACCTGCAGTATTT 12099
QY 2629 TGGAAAGCAA 2639
Db 12100 TGGAAAGCACA 12110

RESULT 4
US-09-764-868-493:
; Sequence 493, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P732
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 493
; LENGTH: 1316
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12H1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1299)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-493:

Query Match 48.1%; Score 1286; DB 9; Length 1316;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1286; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 45 ACGGGACGCGCAGCGCGCTTGCACCCGGGATGGCGAGGCCCGGAGCGTCCCCGCC 104
Db 1 ACGGGACGCGCAGCGCGCTTGCACCCGGGATGGCGAGGCCCGGAGCGTCCCCGCC 60
QY 105 TGCAGTCCGAGCTACGACCTCACGGCAAGGTGATGCTTCGGGAGACACAGGCGTGGC 164
Db 61 TGCAGTCCGAGCTACGACCTCACGGCAAGGTGATGCTTCGGGAGACACAGGCGTGGC 120
QY 165 AAACATGTTTCTGTATCCATTCAAAGACGGGCTTCTCTCGGAACTTCATAGCC 224
Db 121 AAACATGTTTCTGTATCCATTCAAAGACGGGCTTCTCTCGGAACTTCATAGCC 180
QY 225 ACCGTCCGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGGAGAGTGAAGCTG 284
Db 181 ACCGTCCGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGGAGAGTGAAGCTG 240
QY 285 CAGATCTGGGACACCGCTGGGACGAAAGGTTCCGAAAGCGTCACCCATGCTTATACAGA 344
Db 241 CAGATCTGGGACACCGCTGGGACGAAAGGTTCCGAAAGCGTCACCCATGCTTATACAGA 300
QY 345 GATCTCAGGCGCTTCTCTCTGTATGACATACCAACAAATCTTCTTCGACCAATC 404
Db 345 GATCTCAGGCGCTTCTCTCTGTATGACATACCAACAAATCTTCTTCGACCAATC 404
```

```
Db 301 GATGCTCAGGCTTCTCTCTGTATGACATCACCACAAATCTTCTTCGACAAATC 360
QY 405 AGGGCTTGGCTACTGAGATTATGAGTATCCAGAGGAGCGTGGTGTATGATGCTGCTA 464
Db 361 AGGGCTTGGCTACTGAGATTATGAGTATCCAGAGGAGCGTGGTGTATGATGCTGCTA 420
QY 465 GGCACAGAGGGGATATGAGCAGGAAAGAGTATCCCTTCGAGAGCGGAGACCTTG 524
Db 421 GGCACAGAGGGGATATGAGCAGGAAAGAGTATCCCTTCGAGAGCGGAGACCTTG 480
QY 525 GCCAGGAGTAGTGGTGTCTCTCTGAGACCGACGACCAAGACTGGCATGAATGTGGAG 584
Db 481 GCCAGGAGTAGTGGTGTCTCTCTGAGACCGACGACCAAGACTGGCATGAATGTGGAG 540
QY 595 TTAGCTTTCTGGCCATCGCCAAAGAACTGAAATACCGGGCCGGGCATCAGCGGATGAG 644
Db 541 TTAGCTTTCTGGCCATCGCCAAAGAACTGAAATACCGGGCCGGGCATCAGCGGATGAG 600
QY 645 CCCAGCTTCCAGATCCGAGACTATGAGTGTCCAGAGTCCCAAGAGAGCGTCCAGCTGTCTCC 704
Db 601 CCCAGCTTCCAGATCCGAGACTATGAGTGTCCAGAGTCCCAAGAGAGCGTCCAGCTGTCTCC 660
QY 705 TTATGTGAATCCAGGGGCGAGAGAGGAGCTCTGGAGGCACACAGGATCAGCTTCC 764
Db 661 TTATGTGAATCCAGGGGCGAGAGAGGAGCTCTGGAGGCACACAGGATCAGCTTCC 720
QY 765 CCCTCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGAGGAGTCC 824
Db 721 CCCTCCAGGCTGGCTTATTCAGAGGCTGAGCCAAATGGGAGAAAGATGAGGAGTCC 780
QY 825 ACTGACAGCGCTTCTAGCAGGAGCTTACTTACTTACTTACTTACTTACTTACTTACTT 884
Db 781 ACTGACAGCGCTTCTAGCAGGAGCTTACTTACTTACTTACTTACTTACTTACTTACTT 840
QY 885 CTCCTCCGATCCAGGAGGCTTAAACACTTACTTACTTACTTACTTACTTACTTACTTACTT 944
Db 841 CTCCTCCGATCCAGGAGGCTTAAACACTTACTTACTTACTTACTTACTTACTTACTTACTT 900
QY 945 ATACCAAAAGGCGCTGGATCCCAAAAGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 1004
Db 901 ATACCAAAAGGCGCTGGATCCCAAAAGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 960
QY 1005 CTTTCTAGGACTTGGGGGGCGGCTCCCTTCTTAAAGCAAAAGTGGTGTCTTCC 1064
Db 961 CTTTCTAGGACTTGGGGGGCGGCTCCCTTCTTAAAGCAAAAGTGGTGTCTTCC 1020
QY 1065 AGCTCAGCCCAAGGGGACACAGATGCTTGGGGGTGAGGCGAGGTAATGACTCCATCG 1124
Db 1021 AGCTCAGCCCAAGGGGACACAGATGCTTGGGGGTGAGGCGAGGTAATGACTCCATCG 1080
QY 1125 CACCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCGAGCTTCACTGTCTCCCGCTCT 1184
Db 1081 CACCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCGAGCTTCACTGTCTCCCGCTCT 1140
QY 1185 CCAGGAGCTTATCTTCCGCCCATCTCCCAATAAGTGGGCCCTTGTCTGTGAGGAAGAC 1244
Db 1141 CCAGGAGCTTATCTTCCGCCCATCTCCCAATAAGTGGGCCCTTGTCTGTGAGGAAGAC 1200
QY 1245 CAAAGCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGAGGACAAAGGGCGAGAG 1304
Db 1201 CAAAGCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGAGGACAAAGGGCGAGAG 1260
QY 1305 TAGGGTCTAGCTGTATCTCTGCGCTTACTAACCACCCCGCTGGAG 1350
Db 1261 TAGGGTCTAGCTGTATCTCTGCGCTTACTAACCACCCCGCTGGAG 1306
```

## RESULT 5

; Sequence 959, Application US/10037270  
; Publication No. US20030104529A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungling  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Tillinghast, John  
; APPLICANT: Drmanac, Radolje T.  
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/10/037,270  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 959  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (124)..(699)  
US-10-037-270-559

Query Match 41.1%; Score 1099.6; DB 9; Length 1106;  
Best Local Similarity 99.6%; Pred. No. 7e-295;  
Matches 1102; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 15 GGCCTGCTCACCCTCTGTCAGGAGACATGACGGGCGGAGCGGCGGCGGCGGCGGCGG 74
Db 1 GGAAGTCTCACCCTCTGTCAGGAGACATGACGGGCGGAGCGGCGGCGGCGGCGGCGG 60
QY 75 GATGGGAGGCGGCGGAGCGCTCCCGCCCTGCACTGAGCTGAGCTTACGACCTCAGCGGCAAG 134
Db 61 GATGGGAGGCGGCGGAGCGCTCCCGCCCTGCACTGAGCTGAGCTTACGACCTCAGCGGCAAG 120
QY 135 GTGATGCTTCTGGGAGACACAGGCGTGGGCAAAACATGTTTCTGATCCAAATTCAGAGAC 194
Db 121 GTGATGCTTCTGGGAGACACAGGCGTGGGCAAAACATGTTTCTGATCCAAATTCAGAGAC 180
QY 195 GGGGCTTCTCTGTCGGGAACTTATAGCCACCGTCCGCAATAGACTTCCAGGAACAGGTG 254
Db 181 GGGGCTTCTCTGTCGGGAACTTATAGCCACCGTCCGCAATAGACTTCCAGGAACAGGTG 240
QY 255 GTGACTGTGAGTGGCGTGGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGGAGCG 314
Db 241 GTGACTGTGAGTGGCGTGGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGGAGCG 300
QY 315 TTCCGAGGCGTCAACCCATCTTATACAGATGCTTACGAGATGCTTACGAGCTTCTGCTGTATGAC 374
Db 301 TTCCGAGGCGTCAACCCATCTTATACAGATGCTTACGAGATGCTTACGAGCTTCTGCTGTATGAC 360
QY 375 ATACCAACAAATCTTCTTCGACAACTAGGCGCTGGCTCAGTCTGAGATTCATGAGTAT 434
Db 361 ATACCAACAAATCTTCTTCGACAACTAGGCGCTGGCTCAGTCTGAGATTCATGAGTAT 420
QY 435 GCCCAGAGGAGCGTGGTGTATCATCTGCTAGGCAACAGGCGGATATGAGCAGGGAAGA 494
Db 421 GCCCAGAGGAGCGTGGTGTATCATCTGCTAGGCAACAGGCGGATATGAGCAGGGAAGA 480
QY 495 GTGATCGTTCGGAAGAGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCTCTCTGGAG 554
```

Db 481 GTGATCCGTTCCGAAGCAGGAGAGACTTGGCCAGGGAGTACGGTTCCTTCCCTGGAG 540  
QY 555 ACCAGGCCAAGACTGGCATGATGTGGAGTTAGCCTTTCTGGCCATCCCAAGGAAGT 614  
Db 541 ACCAGGCCAAGACTGGCATGATGTGGAGTTAGCCTTTCTGGCCATCCCAAGGAAGT 600  
QY 615 AATATCCGGGCGGGGATCAGGGGATGAGCCAGTTCAGATCCGAGACTATGTAGAG 674  
Db 601 AATATCCGGGCGGGGATCAGGGGATGAGCCAGTTCAGATCCGAGACTATGTAGAG 660  
QY 675 TCCAGAAGAAGCGCTCCAGCTGCTCCTTCATGTGAATCCAGGGGCGAGAGAG 734  
Db 661 TCCAGAAGAAGCGCTCCAGCTGCTCCTTCATGTGAATCCAGGGGCGAGAGAG 720  
QY 735 GCTCTGGAGGCACACAGGATGAGCCTTCCCTCCAGGCGCTGGCTTATTCAGAGGC 794  
Db 721 GCTCTGGAGGCACACAGGATGAGCCTTCCCTCCAGGCGCTGGCTTATTCAGAGGC 780  
QY 795 TGAGCCATGGGGAGAAAGATGGAGACTCAGTGCACAGCGCTTCTAGCAGGAGCTA 854  
Db 781 TGAGCCATGGGGAGAAAGATGGAGACTCAGTGCACAGCGCTTCTAGCAGGAGCTA 840  
QY 855 TACTCAACTCCTACTTGTAGTTCTCGGCTTCCCGCATCCACAGGAGGTAACAC 914  
Db 841 TACTCAACTCCTACTTGTAGTTCTCGGCTTCCCGCATCCACAGGAGGTAACAC 900  
QY 915 TTAGCTTTTATTTAATAGTACATAATTAATACCAAAAAAGCGCTGGATCCCCAAAA 974  
Db 901 TTAGCTTTTATTTAATAGTACATAATTAATACCAAAAAAGCGCTGGATCCCCAAAA 960  
QY 975 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTGGGGGCGGCCCTCC 1034  
Db 961 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTGGGGGCGGCCCTCC 1020  
QY 1035 TCCTAAGCATACAAAGGTGGTGTCTCCAGCTCAGCCAGGCGNACAGATGCACCT 1094  
Db 1021 TCCTAAGCATACAAAGGTGGTGTCTCCAGCTCAGCCAGGCGNACAGATGCACCT 1080  
QY 1095 TGGGGGTGAGGCGAGGTAATGACTCC 1120  
Db 1081 TGGGGGTGAGGCGAGGTAATGACTCC 1106

## RESULT 6

US-09-794-257-13  
; Sequence 13, Application US/09794257  
; Patent No. US2002009804A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 32705, 27423, 32700, 32712, No. US2002009804A1el  
; FILE OF INVENTION: Human G-Proteins  
; FILE REFERENCE: 35800/209285  
; CURRENT APPLICATION NUMBER: US/09/794, 257  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/185,606  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (124)...(699)  
US-09-794-257-13

Query Match 40.6%; Score 1086; DB 10; Length 1116;

Best Local Similarity 99.5%; Pred. No. 4.2e-291;

Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGACATAGAGGGCAGCCAGGCGCGGTGGCCACCCGGGATGCGAGGCCCCCGAGGGCT 96

Db 23 GCCTCGGACGGGCACGCCAGGCGCGTTGCCACCGGGATGCGAGGCCCGCCAGCGCT 82  
QY 97 CCCGCGCTTCAGTCCGAGCTACGACTCACGGCAAGGTGATGCTTCTGGGAGACACAG 156  
Db 83 CCCGCGCTTCAGTCCGAGCTACGACTCACGGCAAGGTGATGCTTCTGGGAGACACAG 142  
QY 157 GCGTCGGCAAAACATGTTCTCTGATCCAAATCAAGACAGGGGCTTCTCTCGGAACT 216  
Db 143 GCGTCGGCAAAACATGTTCTCTGATCCAAATCAAGACAGGGGCTTCTCTCGGAACT 202  
QY 217 TCATAGCCACCGTCGCATAGACTTCAGGAACAAGTGGTGAAGTGGATGGCGTGAGAG 276  
Db 203 TCATAGCCACCGTCGCATAGACTTCAGGAACAAGTGGTGAAGTGGATGGCGTGAGAG 262  
QY 277 TGAAGTTCAGATCTGGGACACCGCTGGGACGAGCGTTCCGAAGCGTCCACCATGCTT 336  
Db 263 TGAAGTTCAGATCTGGGACACCGCTGGGACGAGCGTTCCGAAGCGTCCACCATGCTT 322  
QY 337 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCAACAACAAATCTTCTTCG 396  
Db 323 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCAACAACAAATCTTCTTCG 382  
QY 397 ACACATCAGGCGCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 456  
Db 383 ACACATCAGGCGCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 442  
QY 457 TGCCTGTAGCAACAAAGCGGATATGAGACGCAAGAGTATGATGATGATGATGATGATG 516  
Db 443 TGCCTGTAGCAACAAAGCGGATATGAGACGCAAGAGTATGATGATGATGATGATG 502  
QY 517 AGACCTTGGCGAGGAGTACGGTGTTCCTTCTCTGGAGACCGAGCGCAAGACTGGCATGA 576  
Db 503 AGACCTTGGCGAGGAGTACGGTGTTCCTTCTCTGGAGACCGAGCGCAAGACTGGCATGA 562  
QY 577 ATGTGAGTGTAGCTTCTTGGCCATCCCAAGNACTGAATACCGGCGCGGCGCATCAGG 636  
Db 563 ATGTGAGTGTAGCTTCTTGGCCATCCCAAGNACTGAATACCGGCGCGGCGCATCAGG 622  
QY 637 CGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGTCCAGCT 696  
Db 623 CGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGTCCAGCT 682  
QY 697 GCTGCTCTTCATGTGATCCAGGGGCGAGAGGAGGCTGTGGAGGCGACACAGGATGC 756  
Db 683 GCTGCTCTTCATGTGATCCAGGGGCGAGAGGAGGCTGTGGAGGCGACACAGGATGC 742  
QY 757 AGCTTCCCGCTCCAGGCGCTGCTTATCCAAAGAGCTGAGCCAAATGGGAGAAAGATG 816  
Db 743 AGCTTCCCGCTCCAGGCGCTGCTTATCCAAAGAGCTGAGCCAAATGGGAGAAAGATG 802  
QY 817 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATACTCCAAGTCTTACTTGAAGT 876  
Db 803 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATACTCCAAGTCTTACTTGAAGT 862  
QY 877 CTTGCGGTCTCCCGCATCCAGGAGGAGTAAACACTTAGCTTTTATTTAATAGTAC 936  
Db 863 CTTGCGGTCTCCCGCATCCAGGAGGAGTAAACACTTAGCTTTTATTTAATAGTAC 922  
QY 937 ATATTTAATACCAAAAAGCGCGCTGGATCCCCAAAACCGAGGCTGGAGCTAGTGG 996  
Db 923 ATATTTAATACCAAAAAGCGCGCTGGATCCCCAAAACCGAGGCTGGAGCTAGTGG 982  
QY 997 CCCTTTTGTCTTAGGACTTGGGGCGCGCGCTTCTCTTAAAGCATAAACAAAGTGT 1056  
Db 983 CCCTTTTGTCTTAGGACTTGGGGCGCGCGCTTCTCTTAAAGCATAAACAAAGTGT 1042  
QY 1057 GTTCTCCAGCTCAGCCCCAGGGGACACAGATGCATCTTTGGGGGTGAGGCGAGTATGA 1116  
Db 1043 GTTCTCCAGCTCAGCCCCAGGGGACACAGATGCATCTTTGGGGGTGAGGCGAGTATGA 1102  
QY 1117 CTGATGCGACCT 1130  
|||||

Db 1103 CTCCATCGCACCT 1116

## RESULT 7

US-10-051-986-10

; Sequence 10, Application US/10051986

; Patent No. US20020146770A1

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; Tang, Y. Tom

; Lal, Preeti

; Guegler, Karl J.

; Corley, Neil C.

; Patterson, Chandra

; Batra, Sajeev R.

; Baughn, Mariah R.

; TITLE OF INVENTION: RAS PROTEINS

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: US

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/051,986

; FILING DATE: 15-Jan-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/766,551

; FILING DATE: DECEMBER 12, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Cerrone, Michael C.

; REGISTRATION NUMBER: 39,132

; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 875 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: UCMCL5T01

; CLONE: 1528559

; SEQUENCE DESCRIPTION: SEQ ID NO: 10 :

US-10-051-986-10

Query Match 32.6%; Score 873; DB 12; Length 875;

Best Local Similarity 100.0%; Pred. No. 5.1e-232;

Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 CGGCGACCCAGCGCGTTCGCCACCGGGATGGGAGGCCCCCGAGGGCTCCCGCGCCT 105

DB 1 CGGGCAGCCAGCGCGCTTCGCCACCGGGATGGGAGGCCCCCGAGGGCTCCCGCGCCT 60

QY 106 GCAGTCGAGCTAGCACTACGGCGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGCA 165

DB 61 GCAGTCGAGCTAGCACTACGGCGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGCA 120

QY 166 AAACATGTTTCTGATCCAAATTAAGACGGGGCTTCCTGTCGGGAACCTTCATAGCCA 225

DB 121 AAACATGTTTCTGATCCAAATTAAGACGGGGCTTCCTGTCGGGAACCTTCATAGCCA 180

QY 226 CCCTCGGCATAGACTTCAGGAACAAGGTGGTGAATGGCCGTGAGTGAAGTGC 285  
DB 181 CCCTCGGCATAGACTTCAGGAACAAGGTGGTGAATGGCCGTGAGTGAAGTGC 240  
QY 286 AGATCTGGACACCGCTGGGAGGAAACGGTTCGGAAGCGTCACCCATCTTATTACAGAG 345  
DB 241 AGATCTGGACACCGCTGGGAGGAAACGGTTCGGAAGCGTCACCCATCTTATTACAGAG 300  
QY 346 ATGCTCAGGCCTTGCTTCTGCTGATGACATCACCAACAAATCTCTTCGACACATCA 405  
DB 301 ATGCTCAGGCCTTGCTTCTGCTGATGACATCACCAACAAATCTCTTCGACACATCA 360  
QY 406 GGGCTGCTCACTGAGATTTCATGATGATGCCAGAGGAGCGTGTGATCATCTGCTAG 465  
DB 361 GGGCTGCTCACTGAGATTTCATGATGATGCCAGAGGAGCGTGTGATCATCTGCTAG 420  
QY 466 GCAACAAGCGGATATGACAGCGAAAGAGTATCCGTTCCGAACACGAGAGACCTTGG 525  
DB 421 GCAACAAGCGGATATGACAGCGAAAGAGTATCCGTTCCGAACACGAGAGACCTTGG 480  
QY 526 CCAGGAGTACGGTGTTCCTTCTGAGACCGCCAGGACTGGCATGATGTGGAGT 585  
DB 481 CCAGGAGTACGGTGTTCCTTCTGAGACCGCCAGGACTGGCATGATGTGGAGT 540  
QY 586 TAGCCTTCTGCGCATCCCAAGGAACTGAATACCGCGCGCGGCGGATGAGC 645  
DB 541 TAGCCTTCTGCGCATCCCAAGGAACTGAATACCGCGCGCGGCGGATGAGC 600  
QY 646 CCAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAGCGCTCCAGCTGCTCCT 705  
DB 601 CCAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAGCGCTCCAGCTGCTCCT 660  
QY 706 TCATGTGAATCCAGGGGCGAGAGAGGCTCTTGAGGACACACAGGATGAGCCTTCCC 765  
DB 661 TCATGTGAATCCAGGGGCGAGAGAGGCTCTTGAGGACACACAGGATGAGCCTTCCC 720  
QY 766 CCTCCAGGCTGCGCTTATTCACAGAGGCTGAGCCATGGGAGAAAGATGAGGACTCA 825  
DB 721 CCTCCAGGCTGCGCTTATTCACAGAGGCTGAGCCATGGGAGAAAGATGAGGACTCA 780  
QY 826 CTGCACAGCGCTTCCTAGCAGGAGCTATATCTCAACTCTTACTTGTGTTCTCGGTC 885  
DB 781 CTGCACAGCGCTTCCTAGCAGGAGCTATATCTCAACTCTTACTTGTGTTCTCGGTC 840  
QY 886 TCCCGCATCCACAGGAGGCTAAACACTTAG 918  
DB 841 TCCCGCATCCACAGGAGGCTAAACACTTAG 873

## RESULT 8

US-10-274-968-3

; Sequence 3, Application US/10274968

; Publication No. US20030049793A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER

; FILE REFERENCE: CL001052 DIV

; CURRENT APPLICATION NUMBER: US/10/274,968

; PRIOR FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/740,027

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 24707

; TYPE: DNA

; ORGANISM: Human

; US-10-274-968-3

Query Match

Best Local Similarity 25.6%; Score 683.4; DB 9; Length 24707;

Matches 873; Conservative 99.9%; Pred. No. 1.4e-178;

|    | Matches | 684;   | Conservative | 0; | Mismatches | 1; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|----|--------|----|------|----|
| Qy | 1955    | GTCTTGTGAAAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGGAGACAAGATTAAAGC     | 2014         |    |            |    |        |    |      |    |
| Db | 1       | GTCTTGTGAAAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGGAGACAAGATTAAAGC     | 60           |    |            |    |        |    |      |    |
| Qy | 2015    | AAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTCGACACCCCAT       | 2074         |    |            |    |        |    |      |    |
| Db | 61      | AAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTCGACACCCCAT       | 120          |    |            |    |        |    |      |    |
| Qy | 2075    | GCCTTTTCCCTTTAACTTCTCAACACAGATACCAAGGGCTTAAACTGCTTTACTCTCCCTCCCT | 2134         |    |            |    |        |    |      |    |
| Db | 121     | GCCTTTTCCCTTTAACTTCTCAACACAGATACCAAGGGCTTAAACTGCTTTACTCTCCCTCCCT | 180          |    |            |    |        |    |      |    |
| Qy | 2135    | ACTGAGTCAGGTTAGTGGTGGGAGGTCACCCCATTTCCGAGTTTAAACCAATGCAATATGA    | 2194         |    |            |    |        |    |      |    |
| Db | 181     | ACTGAGTCAGGTTAGTGGTGGGAGGTCACCCCATTTCCGAGTTTAAACCAATGCAATATGA    | 240          |    |            |    |        |    |      |    |
| Qy | 2195    | GTAACCAAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTC    | 2254         |    |            |    |        |    |      |    |
| Db | 241     | GTAACCAAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTC    | 300          |    |            |    |        |    |      |    |
| Qy | 2255    | CTCCTTGGTCACATATCTCCCAAGCTCTGATCCCTGGCCATGGGAAGTGCACAGGAACA      | 2314         |    |            |    |        |    |      |    |
| Db | 301     | CTCCTTGGTCACATATCTCCCAAGCTCTGATCCCTGGCCATGGGAAGTGCACAGGAACA      | 360          |    |            |    |        |    |      |    |
| Qy | 2315    | TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCTTGAGGGGGGAGAGGGG     | 2374         |    |            |    |        |    |      |    |
| Db | 361     | TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCTTGAGGGGGGAGAGGGG     | 420          |    |            |    |        |    |      |    |
| Qy | 2375    | GAGGAAGAGTATGCGGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCGAAT        | 2434         |    |            |    |        |    |      |    |
| Db | 421     | GAGGAAGAGTATGCGGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCGAAT        | 480          |    |            |    |        |    |      |    |
| Qy | 2435    | CTTGCTCTGAGCAGTCAGCGGGCTCCAGTTTGGCCGATGAAGGAAGTTCTCCCTGGCCCTC    | 2494         |    |            |    |        |    |      |    |
| Db | 481     | CTTGCTCTGAGCAGTCAGCGGGCTCCAGTTTGGCCGATGAAGGAAGTTCTCCCTGGCCCTC    | 540          |    |            |    |        |    |      |    |
| Qy | 2495    | CCTCAGGCAGAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGGCCCAAGGCAGGT       | 2554         |    |            |    |        |    |      |    |
| Db | 541     | CCTCAGGCAGAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGGCCCAAGGCAGGT       | 600          |    |            |    |        |    |      |    |
| Qy | 2555    | TGCAGGAGATCCAAATCCCATAGACAGCTCTGGGCCCTTCGATTTGAGTTTTTCAGNAATT    | 2614         |    |            |    |        |    |      |    |
| Db | 601     | TGCAGGAGATCCAAATCCCATAGACAGCTCTGGGCCCTTCGATTTGAGTTTTTCAGNAATT    | 660          |    |            |    |        |    |      |    |
| Qy | 2615    | AAACTGCAGTATTTTGGAAAGCAAA  | 2639         |    |            |    |        |    |      |    |
| Db | 661     | AAACTGCAGTATTTTGGAAAGCAAA  | 685          |    |            |    |        |    |      |    |

```

RESULT 9
US-09-740-027-3
: Sequence 3, Application US/09740027
: Patent No. US20020076749A1
: GENERAL INFORMATION:
: APPLICANT: YE, Jane et al
: TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER
: TITLE OF INVENTION: COFACTOR PROTEINS, AND USES THEREOF
: FILE REFERENCE: CL001052
: CURRENT APPLICATION NUMBER: US/09/740,027
: CURRENT FILING DATE: 2000-12-20
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 24707
: TYPE: DNA
: ORGANISM: Human
: US-09-740-027-3

Query Match          25.6%   Score 693.4; DB 10; Length 24707;
Best Local Similarity 99.9%   Pred. No. 1.4e-178;

```

Query Match 25.6%; Score 683.4; DB 10; Length 24707;  
Best Local Similarity 99.9%; Pred. No. 1.4e-178;

| Matches | 684; | Conservative   | 0;   | Mismatches | 1; | Indels | 0; | Gaps | 0; |
|---------|------|--|------|------------|----|--------|----|------|----|
| Qy      | 1955 | GTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGGAGACAAGATTAAAGC   | 2014 |            |    |        |    |      |    |
| Db      | 1    | GTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGGGAGACAAGATTAAAGC   | 60   |            |    |        |    |      |    |
| Qy      | 2015 | AAAGCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCTCTCCCTGACCACCCAT   | 2074 |            |    |        |    |      |    |
| Db      | 61   | AAAGCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCTCTCCCTGACCACCCAT   | 120  |            |    |        |    |      |    |
| Qy      | 2075 | GCCTTTCCCTTTAACTTCTCAACACAGATACGAGGGCTTAAACTGCTTTACTCTCCCTCCCT | 2134 |            |    |        |    |      |    |
| Db      | 121  | GCCTTTCCCTTTAACTTCTCAACACAGATACGAGGGCTTAACTGCTTTACTCTCCCTCCCT  | 180  |            |    |        |    |      |    |
| Qy      | 2135 | ACTGAGTCAGGTTAGTGGTGGGAGGTCACCCATTTCCGAGTTTAAACCAATGCAATATGA   | 2194 |            |    |        |    |      |    |
| Db      | 181  | ACTGAGTCAGGTTAGTGGTGGGAGGTCACCCATTTCCGAGTTTAAACCAATGCAATATGA   | 240  |            |    |        |    |      |    |
| Qy      | 2195 | GTAACAAAGTCATGTGGGTATGCTCGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC      | 2254 |            |    |        |    |      |    |
| Db      | 241  | GTAACAAAGTCATGTGGGTATGCTCGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC      | 300  |            |    |        |    |      |    |
| Qy      | 2255 | CTCCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGGAAGTGACAGGAAACA     | 2314 |            |    |        |    |      |    |
| Db      | 301  | CTCCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGGAAGTGACAGGAAACA     | 360  |            |    |        |    |      |    |
| Qy      | 2315 | TGAGGTCATCACCTGCAGGCATCTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGAGGG     | 2374 |            |    |        |    |      |    |
| Db      | 361  | TGAGGTCATCACCTGCAGGCATCTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGAGGG     | 420  |            |    |        |    |      |    |
| Qy      | 2375 | GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT   | 2434 |            |    |        |    |      |    |
| Db      | 421  | GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT   | 480  |            |    |        |    |      |    |
| Qy      | 2435 | CTTGCTCTGACAGTCAGCGCTCCAGTTTGGGCCCATTAAGGAAGTTCCTCGTGGCCTC     | 2494 |            |    |        |    |      |    |
| Db      | 481  | CTTGCTCTGACAGTCAGCGCTCCAGTTTGGGCCCATTAAGGAAGTTCCTCGTGGCCTC     | 540  |            |    |        |    |      |    |
| Qy      | 2495 | CCTCAGGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTCTCGGGGCCCAAGGCAGGT     | 2554 |            |    |        |    |      |    |
| Db      | 541  | CCTCAGGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTCTCGGGGCCCAAGGCAGGT     | 600  |            |    |        |    |      |    |
| Qy      | 2555 | TGCAGGAGATCCCATCCATAGACAGCTCTGGGCCCTCTTGCATTTGAGTTTTTCAGAAAT   | 2614 |            |    |        |    |      |    |
| Db      | 601  | TGCAGGAGATCCCATCCATAGACAGCTCTGGGCCCTCTTGCATTTGAGTTTTTCAGAAAT   | 660  |            |    |        |    |      |    |
| Qy      | 2615 | AAACTGCAGTATTTTGGAAAGCAAA                                      | 2639 |            |    |        |    |      |    |
| Db      | 661  | AAACTGCAGTATTTTGGAAAGCAAA                                      | 685  |            |    |        |    |      |    |

```

RESULT 10
US-09-817-199A-24
: Sequence 24, Application US/09817199A
: Patent No. US20020142380A1
: GENERAL INFORMATION:
: APPLICANT: SHAO, Wei et al.
: TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
: TITLE OF INVENTION: PROTEINS, AND USES THEREOF
: FILE REFERENCE: CL001187
: CURRENT APPLICATION NUMBER: US/09/817,199A
: CURRENT FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 24
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-817-199A-24

Query Match          22.5%; Score 600.6; DB 10; Length 601;
Best Local Similarity 99.8%; Pred. No. 1.8e-156;

```

Query Match . 22.5%; Score 600.6; DB 10; Length 601;  
Best Local Similarity 99.8%; Pred. No. 1.8e-156;





